seit 1558

Small noise analysis of time-periodic bistable jump diffusions

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

vorgelegt dem Rat der

Fakultät für Mathematik und Informatik der

Friedrich-Schiller-Universität Jena

von Diplom-Mathematikerin Isabelle Kuhwald

geboren am 21.09.1986 in Jena

Gutachter:

1. Prof. Dr. Ilya Pavlyukevich, Friedrich-Schiller-Universität Jena
2. Prof. Dr. Samuel Herrmann, Université de Bourgogne (Dijon)

Tag der mündlichen Prüfung: 19. Juni 2015

Zusammenfassung

Das kontraintuitive Phänomen der Stochastischen Resonanz beschreibt das Verstärken schwacher periodischer Eingangssignale von nicht linearen Systemen durch Hinzufügen von Rauschtermen mit geringer Amplitude. In Systemen, die Gaußschem Rauschen ausgesetzt sind, wurde Stochastische Resonanz schon eingehend untersucht. Im Folgenden dienen Sprungprozesse als Rauschterm.

Zu Beginn wird die gestörte Bewegung eines Partikels innerhalb eines Zweitopfpotentials mit periodisch verändernden Minimumpositionen analysiert. Störterme, die Sprünge zulassen und polynomiell fallende Verteilungsschwänze aufweisen, werden genauer betrachtet. Zeitstetige Markov-Ketten mit zwei Zuständen dienen als erste Approximation der Bewegung des Partikels zwischen den zwei Potentialtöpfen. Die Übergangszeiten der Ketten zwischen den zwei Zuständen wird dabei periodisch und polynomiell in der Rauschamplitude gewählt. Die Zeitskala, auf der sich die Markov-Ketten nahezu periodisch verhalten, ist die Inverse der mittleren Übergangszeit. Dies kann durch die Verwendung eines wahrscheinlichkeitstheoretischen Qualitätsmaßes, welches die Wahrscheinlichkeit eines Sprunges innerhalb eines kleinen Zeitfensters um den Zeitpunkt an dem ein Übergang am wahrscheinlichsten ist, bewiesen werden. Im Weiteren folgt die Anwendung dieses Maßes auf die Sprungdiffusion. Zwei verschiedene Störterme, die durch zwei unterschiedliche stochastische Integrationsbegriffe entstehen, werden verwendet. Wird die Sprungdiffusion auf der richtigen Zeitskala betrachtet, ist abgesehen von kleinen Lokalisationsfehlern ein Übergang des Partikels von einem beschränkten Anziehungsgebiet in das andere innerhalb eines kleinen Zeitintervalls um den Zeitpunkt, an dem ein Übergang höchstwahrscheinlich ist, ebenso wahrscheinlich wie für die approximierende Markov-Kette.

Des Weiteren wird das Verhalten eines Partikels innerhalb eines zeit-unabhängigen Zweitopfpotentials untersucht, wobei die Bewegung diesmal durch ein dem stabilen Lévy-Rauschen ähnliche Störung mit periodisch schwankendem Stabilitätsparameter gestört wird. Erneut sind die Übergangszeiten polynomiell in der Rauschamplitude, jedoch variiert diesmal der Exponent periodisch. Es stellt sich heraus, dass der minimale Wert des periodischen Stabilitätsparameters des Rauschterms eine wichtige Rolle spielt, um die Zeitskala anzugeben auf der sich die Sprungdiffusion im Wesentlichen wie eine Markov-Kette mit zwei Zuständen verhält. Das bereits genannte wahrscheinlichkeitstheoretische Qualitätsmaß schlägt eine Zeitskala vor, die etwas größer ist, als die Inverse der polynomiellen Zeit mit dem minimalen Wert des Stabilitätsparameters als Exponent. Die richtig skalierte Sprungdiffusion zeigt metastabiles Verhalten, dass heißt auf der erwähnten kritischen Zeitskala ist die wesentliche Dynamik der Sprungdiffusion von dem Verhalten einer Zeit-stetigen Markov-Kette, die nur die Potentialminima als Werte annimmt und der das Springen nur erlaubt ist, wenn eine Minimumposition des periodischen Stabilitätsparameters erreicht wurde, kaum zu unterscheiden.

Abstract

The counterintuitive phenomenon of stochastic resonance describes an enhancement of weak periodic input signals of nonlinear systems evoked by garnishing it with noise of small amplitude. After extensive examinations of stochastic resonance in systems perturbed by Gaussian noise finally jump processes enter the stage.

At the beginning the perturbed movement of a particle in a double-well potential with periodically changing minimum positions is analysed. Perturbation processes that admit jumps and have distributions with polynomially decaying tails are in the focus of our research. Time-continuous two-state Markov chains serve as the first approximation for the hopping dynamics of the particle between the two wells. The transition rates of the chains from one state to the other vary periodically and are chosen to be polynomial with respect to the noise parameter. The time scale on which the Markov chains behave almost periodically is the inverse of the average transition rate. This is proven with the help of a probabilistic quality measure that maximizes the probability of a jump in a small time window around the time point at which a transition is most likely to occur. The transfer of this result to the complex jump diffusion model follows. Two different perturbation terms arising from two distinct understandings of stochastic integration with respect to jump processes are of interest. On the appropriately chosen time scale apart from small localization errors a transition of the particle from one bounded well to the other one in a small time window including the time point at which a transition is most likely to occur is as probable as in the case of the approximating two-state Markov chain.

Furthermore the behaviour of a particle moving in a time-independent double-well potential subject to stable-like Lévy noise with time-periodic stability index is examined. Again a time-continuous two-state Markov chain serves as a simplification. As above, its transition rates are polynomial with respect to the noise amplitude but have periodically varying exponents. It turns out that the minimal value of the periodic stability index of the noise is decisive to specify the time scale on which the jump diffusion behaves like a two-state Markov chain. The already explained probabilistic quality measure proposes a critical time scale which is slightly longer than the inverse of the polynomial time with the exponent equal to the minimal value of the stability index. The appropriately scaled jump diffusion shows metastable behaviour, i.e. on the critical time scale its effective dynamics are indistinguishable from the behaviour of a time-continuous two-state Markov chain living on the well bottoms and only allowed to jump at time points where the minimal value of the periodic stability index is attained.

Acknowledgement

Several years ago Prof. Dr. Ilya Pavlyukevich was the third person I asked for a position as a PhD student. Although we were strangers for each other he gave me a chance to prove my mathematical knowledge and broaden my horizon in numerous fields. I am grateful for his decision and his support during the last years especially in the last months. It is a great pleasure to participate in his train of thoughts and reap the benefits of them. Without his encouragement and guidance I would not have come this far. By now he has become my top choice of a supervisor.

Furthermore, my thanks are due to all present and past members of the department of mathematics at the FSU Jena, especially to all colleagues of the sector stochastic and statistics. First of all, to Prof. Dr. Michael Neumann who stimulated my probabilistic curiosity in an enormous number of lectures and additionally had a considerable share on my appointment as doctoral candidate. Special thanks go to Prof. Dr. Werner Linde whose lectures are outstanding and who was a great supervisor for my diploma thesis. But also to colleagues and former colleagues like Anne, Johannes, Kai, Nadine and Olli and fellow students like Heidi, Philipp and Toralf who accompanied me and made the last years unforgettable.

Besides all this scientific support the moral comfort of my family and friends is gratefully acknowledged. The endless love of my parents and the understanding and encouragement of my sister made this work possible. After diverse setbacks I always regained new strength after an refreshing evening with all my friends.

The biggest part of energy that was necessary to finish this thesis comes from my husband Daniel who for me is like a solid rock in the waves. He never lost his patience and faith in me. At last, the birth of our little daughter gave us a new purpose in life and me a new inspiration to complete what I had not expected to finish ever.

Contents

Chapter 1

Introduction

Features one usually associates with the term “noise” are chaotic, agitated or disordered which all together not give a very positive impression of this phenomenon. It is hard to imagine that an unwanted disturbance may achieve any good. Actually this attitude is a little biased which becomes apparent after the acquaintance with stochastic resonance. Regardless of whether stochastic resonance appears in a system originating from climatology, electronic engineering, biology or medicine it always pursues the same mechanism. A weak periodic input signal of a nonlinear system is enhanced by adding noise with a correct amplitude. Without the otherwise unwanted perturbation, the signal detection worsens noticeably.

1.1 Historical motivation

To trace the origin of stochastic resonance one has to draw the attention to climatology. The climate of the Earth varies in repeating patterns between warm periods and ice ages with temperature gradient of around 10 K. Weak periodic variations of the Earth’s orbit (eccentricity), its axial tilt (obliquity) and the axial precession better known as Milankovich cycles possess periods of around 100.000, 41.000 and 25.000 years. Although these variations correlate with the climate evolution it seems unreasonable that such weak modulations alone cause fast transitions between the two noticeable different climate regimes. Nicolis ([35]) and Benzi et al. ([3]) discovered an amplifying effect of the weather and ocean circulations on the orbital forcing. Thus small weather fluctuations act as an amplifier and allow rapid changes from warm to cold or vice versa.

In 1969 Budyko ([11]) and Sellers ([43]) simultanously developed a very simple model of the Earth climate. This model which is called energy balance model describes the evolution of the spatially (over longitude and latitude) and annually averaged temperature $X(t)$ on Earth and is well explained in [23]. The evolution of the temperature is assumed to be equal to the difference between incoming E_{in} and outgoing energy E_{out} of the Earth. The solar constant $s(t)$ that contrary to its name is not constant but proportional to the incoming radiative

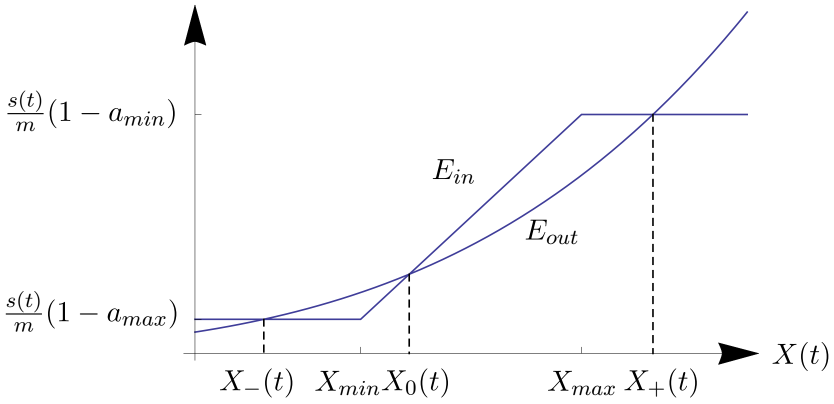


Figure 1.1: The energy balance model

energy is chosen to be periodic to cope with the Milankovich cycles, namely

$$s(t) = s_0 + c\sin\frac{2\pi t}{2T}$$

for special constants s_0 and c and a period $2T$ of around 100.000 years. Additionally E_{in} depends on the ratio of absorbed to incoming radiation. This ratio is given by $1-a$ where a is the so-called albedo that quantifies the proportion of reflected radiation. A snow-covered globe has an essentially bigger albedo than a wooded green and brown Earth. Therefore a temperature below a critical value X_{min} is associated with a big albedo a_{max} and if the temperature lies above X_{max} the albedo obtains its smallest value a_{min}. The course of a is modelled to be linear inbetween. A simple model for E_{out} is the black body radiator. The outgoing energy E_{out} is equated to $\gamma X^4(t)$ for some constant γ. Taking into account that the Earth has heat capacity m yields

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) = \frac{1}{m}\left(E_{in}(t) - E_{out}(t)\right) = \frac{s(t)}{m}\left(1 - a(X(t)\right) - \frac{\gamma}{m}X^4(t).$$

Due to the structure of $E_{in} - E_{out}$ three critical temperature levels $\{X_-(t), X_0(t), X_+(t)\}$ with $0 = E_{in} - E_{out}$ exist (Figure 1.1) which should be distinguished from all the others due to their stability properties. The state $X_-(t)$ is associated with a cold climate and $X_+(t)$ corresponds to a warm period. Both states are stable while $X_0(t)$ represents an unstable state. All points periodically vary in t because of the periodicity of the solar influence.

Since the model is one-dimensional a representation of $\frac{1}{m}(E_{in} - E_{out})$ through the gradient of a potential is always possible. Assume $U\colon \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}_+$ denotes a potential which is 1-periodic in the second variable (time), has two minima located at $m_-(t)$ and $m_+(t)$ and a separating saddle point at $m_0(t) \in (m_-(t), m_+(t))$ (Figure 1.2). The derivative of $X(t)$ can be rewritten by

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t) = -\nabla U\left(X(t), \frac{t}{2T}\right)$$

while $\nabla U(x,t) = \frac{\partial}{\partial x}U(x,t)$. If the starting temperature lies within the well corresponding to

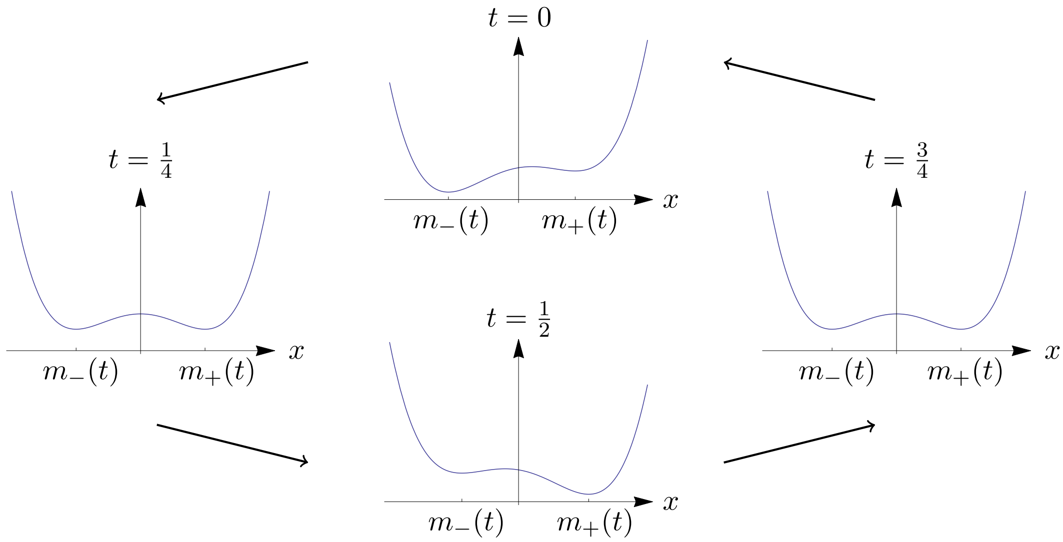


Figure 1.2: The shape of $U(\cdot, t)$ for $t \in \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right\}$.

$m_-(t)$ respectively $m_+(t)$, it is attracted by the region around the well bottom.
The main drawback of this simple deterministic model is the impossibility of climate changes from cold to warm or warm to cold. This difficulty is overcome by adding a noise term $\varepsilon \dot{W}_t$ ([35],[3]), where $\varepsilon > 0$ is small and $\dot{W}_t$ is a white noise:

$$\dot{X}(t) = -\nabla U\left(X(t), \frac{t}{2T}\right) + \varepsilon\, \dot{W}_t. \tag{1.1}$$

Transitions between ice ages and warm periods become possible, but when do they occur? The aim is to detect the underlying periodicity of U in the random output signal. Then the Earth's climate can be truly described by a random enhancement of the weak modulation of the solar constant. Hence the task is to increase the sensitivity of this random system to the periodic modulation of U, which finally justifies the name stochastic resonance. This is done by a careful choice of the noise amplitude ε. Although the addition of $\varepsilon \dot{W}_t$ allows transitions between the wells, excursions into the other well are very rare if ε is very small (Figure 1.3 left). Many opportune moments to jump into the other well will be missed. If ε is chosen very big (Figure 1.3 middle), several transitions can be observed but those numerous excursions cover up the periodic input signal which then cannot be detected. An almost periodic random output signal is only obtained if ε is moderate such that too many transitions are avoided and a stucking within a well is shun (Figure 1.3 right).
Hence it is crucial to tune the noise intensity correctly according to the underlying period of the potential. The rigorous quantification of stochastic resonance is carried out by quality measures of tuning arising either from spectral, information or probalistic theory. These measures precisely synchronize the period $2T$ and the noise amplitude ε to obtain an almost periodic output signal.

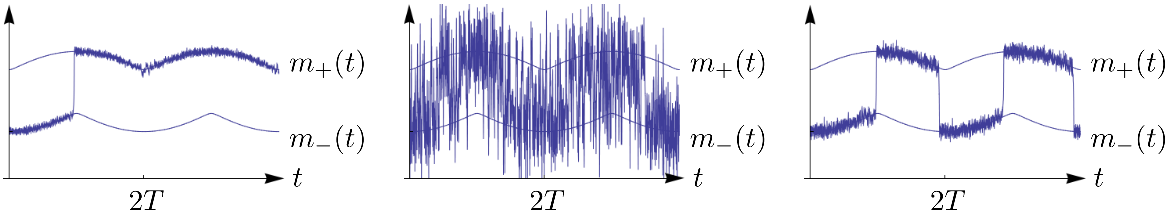


Figure 1.3: Solutions of (1.1) for small, big and moderate values of ε.

1.2 Stochastic resonance and Gaussian processes

One possible approach to stochastic resonance takes into account the consideration of a diffusion that is forced by a gradient system with two attractive points and is subject to noise with small amplitude. Assume Y^ε denotes the solution of the stochastic differential equation

$$\mathrm{d}Y_t^\varepsilon = -\nabla V\left(Y_t^\varepsilon\right)\mathrm{d}t + \varepsilon\,\mathrm{d}W_t, \quad t \geqslant 0$$

where V is a double-well potential, W is a Brownian motion and ε is small. In [16] large deviations theory serves for an analysis of the asymptotic behaviour of Y^ε as ε tends to zero. The process Y^ε on average leaves a well at Kramers' time ([31], [16], [9]) which is of order $\mathrm{e}^{2\Delta V/\varepsilon^2}$ while ΔV denotes the height of the potential barrier.

By adding a periodic modulation to the potential V a time-periodic double well potential U is created and an observation of stochastic resonance is conceivable. Consider the solution $X^{\varepsilon,T}$ of

$$\mathrm{d}X_t^{\varepsilon,T} = -\nabla U\left(X_t^{\varepsilon,T}, \frac{t}{2T}\right)\mathrm{d}t + \varepsilon\,\mathrm{d}W_t, \quad t \geqslant 0 \tag{1.2}$$

while U denotes a double-well potential that is 1-periodic in time and $2T$ is the period length. Now the potential height and therefore the mean exit time from a well periodically vary. Assume the height varies between $\frac{v}{2} > 0$ and $\frac{V}{2} > \frac{v}{2}$. The question is for which $T = T(\varepsilon)$ the periodic behaviour of U transfers to the diffusion $X^{\varepsilon,T}$. Intuitively the half period T must be greater than $\mathrm{e}^{\frac{v}{\varepsilon^2}}$. Then transitions from the shallow into the deep well can be seen. Additionally it seems meaningful to choose T smaller than $\mathrm{e}^{\frac{V}{\varepsilon^2}}$ to avoid exits from the deeper well. With tools from large deviations theory the lower bound was obtained by Freidlin in [15]: The Lebesgue measure of the time the diffusion spends outside a small neighbourhood of the global minimum of U tends to zero in probability if the limit of $\varepsilon \log T(\varepsilon)$ is greater than v.

The qualitative mathematical analysis of stochastic resonance was continued in [37] which was later partly included in the comprehensive work of Herrmann et al. ([20]). In [37] the drift term of the stochastic differential equation in (1.2) uses a potential U which is time-independent on every time interval $[nT,(n+1)T)$, $n \in \mathbb{N}_0$, as seen in Figure 1.4. Two-state Markov chains with periodic infinitesimal generator serve as the first approximation of the

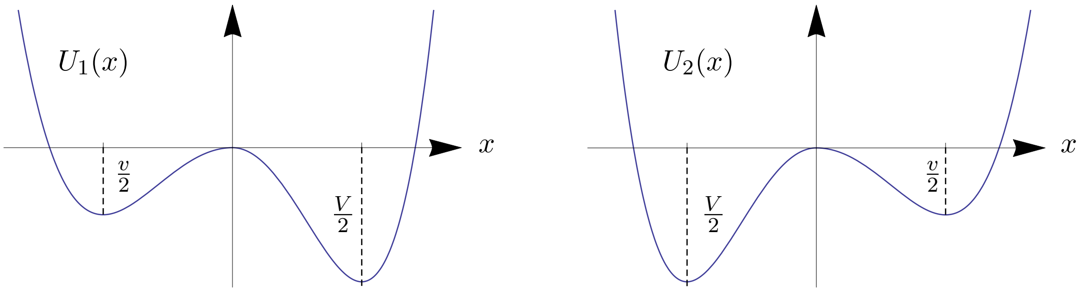


Figure 1.4: In [37] U equals to U_1 on $[0,T),[2T,3T),\ldots$ and $U=U_2$ on $[T,2T),[3T,4T),\ldots$.

essential behaviour of the diffusion $X^{\varepsilon,T}$. In [37] and [24], among others continuous-time Markov chains with two states corresponding to the two minima of U and transition rates equal to the inverse of the mean exit times of the diffusion of one well were considered. The process $X^{\varepsilon,T}$ on average leaves a well at times which are exponentially big in ε. Due to the assymmetry of U two critical exponential time scales $\mathrm{e}^{\frac{v}{\varepsilon^2}}$ and $\mathrm{e}^{\frac{V}{\varepsilon^2}}$ exist and are noticeably separated which is decisive for stochastic resonance.

Many different quality measures of stochastic resonance are applied in [37] to determine the exact value of the half period T that guarantees an almost periodic course of the approximating Markov chain and the diffusion itself. A frequently used tool to detect underlying periodicities is the spectral power amplification coefficient. This measure quantifies the spectral energy at the frequency $\frac{1}{2T}$ of the averaged random output signal. A maximum of the spectral power amplification signals an optimal choice of the pair $2T$ and ε for an at most periodic course of $X^{\varepsilon,T}$ respectively its approximation. While the spectral power amplification of the Markov chain attains its maximum at $T\approx \mathrm{e}^{\frac{(v+V)}{2\varepsilon^2}}$ the quality measure of the diffusion is monotone around this value. This unfavourable behaviour is entitled as lack of robustness and attributes to intra-well fluctuations of the diffusion. A truncation of these fluctuations provides an alternative and reveals the same resonance point $T\approx \mathrm{e}^{\frac{(v+V)}{2\varepsilon^2}}$ for $X^{\varepsilon,T}$. The out-of-phase measure which calculates the time spent in the “wrong location” and therefore totally ignores intra-well movements immediately yields concurrent optimal values for diffusion and Markov chain and thus shows robustness.

Berglund and Gentz applied a pathwise approach ([6]) to analyse the behaviour of a periodically perturbed diffusion in the regime of synchronization where transitions between the two wells become most regular. Singular perturbation and probability theory are used to asymptotically specify the optimal values of noise intensity, amplitude and frequency of the periodic modulation to obtain an almost periodic appearance of the random output signal. One result exactly determines a lower bound of the noise amplitude to make periodic transitions possible. An analysis of residence-time and first-passage-time distributions can be found in [5] where Berglund and Gentz verified that those distributions are very similar to the periodically changing exponential ones.

1.3 Motivation for the application of jump processes

Stationarity, independence of increments and self-similarity are among several good reasons to choose a Brownian motion as perturbation term. However, during the last Ice age 110.000 to 12.000 years before present the Earth's temperature faced several very rapid changes from a cold into an intermediate state and a following slower cooling called Dansgaard-Oeschger-events (Figure 1.5).

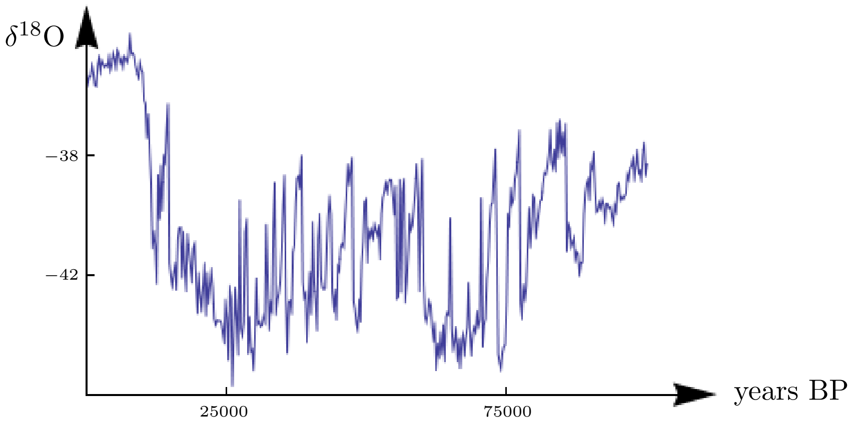


Figure 1.5: An approximation of the Earth temperature from 100.000 years before present until now ([28]) by δ^{18}O values of the NGRIP ice core from Greenland. Low values correspond to low temperatures.

Jump processes that share most of the favourable features of the Brownian motion are α-stable Lévy processes with stability index $\alpha \in (0, 2)$. They can often be considered as a second best choice. While the paths of α-stable processes with small α remind of a step function due to many big jumps, the paths of those processes with α near 2 look more "diffusive" (Figure 1.6).

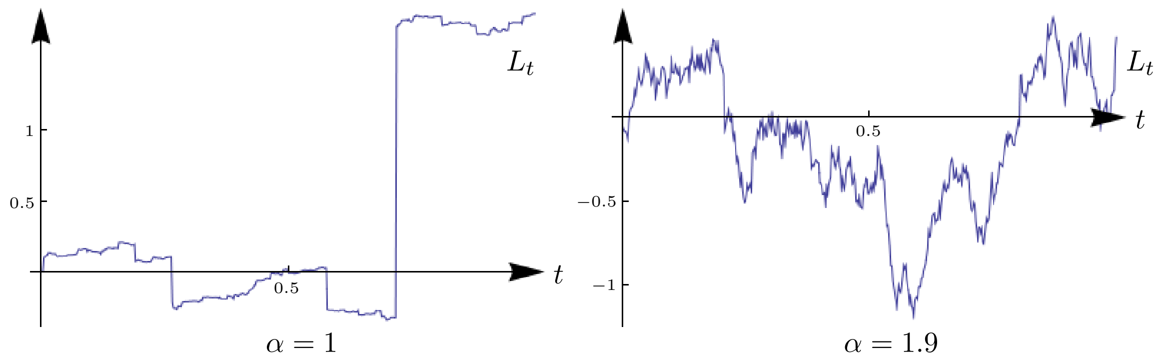


Figure 1.6: Simulation of paths of a symmetric α-stable Lévy process L_t for $\alpha = 1$ (Cauchy process) and $\alpha = 1.9$.

For the first time an energy balance model subject to Lévy flights the physicists term for α-

stable Lévy processes was considered in [13]. Extreme events not only take place in climatic history but also for example at the stock market when a popped bubble causes a crash. Such Lévy processes are important for a variety of models in financial industry ([12], [29]), in chemical physics ([10]), telecommunication ([17]) and many others. An important property of an α-stable Lévy process $L = (L_t)_{t\geqslant 0}$ is the polynomial decay of the tails of its distribution: $\mathbb{P}(|L_1| > x) \approx \frac{c}{x^\alpha}$ as $x \to \infty$ for some $c > 0$. These so-called "heavy tails" might be advantageous over the Brownian motion in models where e.g. crashes, earthquakes, tsunamis and other extreme phenomenona has to be included.

Before we pass to the analysis of stochastic resonance with such jump processes it is appropriate to review results about stochastic differential equations subject to stable Lévy noise. Assume V is a one-dimensional double-well potential with minima at m_- and m_+ and a separating saddle at $m_0 \in (m_-, m_+)$. Consider the solution of the stochastic differential equation

$$\mathrm{d}Y_t^\varepsilon = -\nabla V\left(Y_t^\varepsilon\right)\mathrm{d}t + \varepsilon\,\mathrm{d}L_t, \quad t \geqslant 0.$$

Again the exit times of Y^ε from a well are of interest. The main difference to a Gaussian diffusion is the dependence on the distance to the obstacle in comparison to the barrier height in the case of a Brownian perturbation term which has its origin in the jump affinity of L. The mean exit time from the well corresponding to m_- turns out to be of order $\frac{|m_- - m_0|^\alpha}{\varepsilon^\alpha}$ ([26]). The polynomial order of the mean exit time of the other well is the same which makes a noticeable separation of these critical times impossible. This impedes an observation of stochastic resonance right from the start.

1.4 Overview of this work

The goal of this work is the examination of stochastic resonance for a class of Lévy-driven jump diffusions.

At first we focus on a stochastic differential equation with drift term $U\colon \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}_+$ (see Figure 1.7 for $d = 2$) which is a $\mathbb{R}^d$-dimensional version of the double-well potential seen in Figure 1.2. The perturbation term includes a regularly varying Lévy process L with index $-\alpha < 0$ while regular variation accounts for the heavy tail property. Another generalization is made by considering multiplicative noise terms which allow space dependence in the noise instead of additive ones. The following stochastic differential equation will be studied

$$X_t^\varepsilon(x) = x - \int_0^t \nabla U\left(X_s^\varepsilon, \frac{s}{2T}\right)\mathrm{d}s + \varepsilon\int_0^t g\left(X_{s-}^\varepsilon\right)\mathrm{d}L_s, \quad t \geqslant 0, \tag{1.3}$$

where $\nabla U(x,t) = \left(\frac{\partial}{\partial x_1}U(x,t), \ldots, \frac{\partial}{\partial x_d}U(x,t)\right)$ and the integrand g of the Itô integral in the noise term is a matrix-valued smooth function and the T-dependence of X^ε is omitted for shortness of notation. We even go a step further and devote a part of this thesis to the

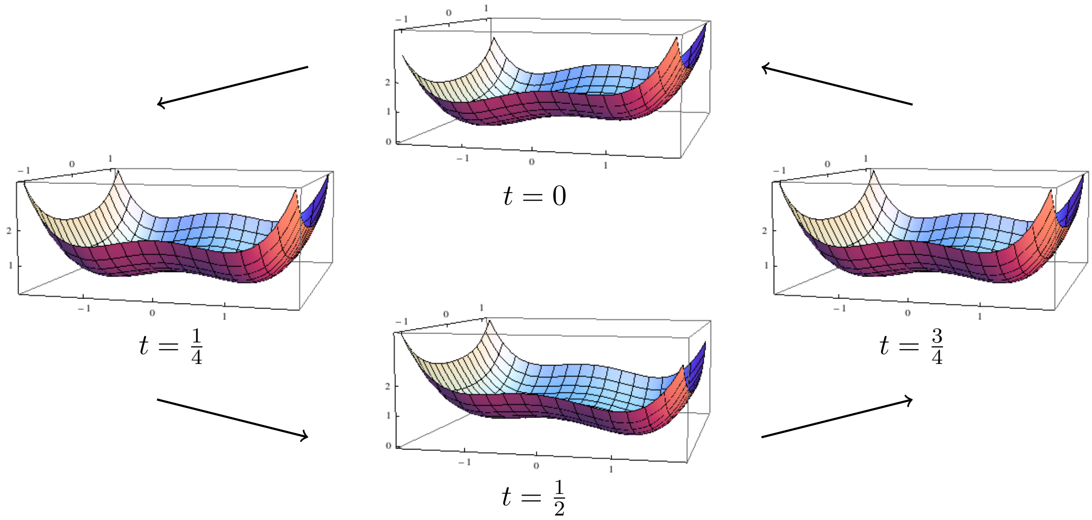


Figure 1.7: If $d = 2$ then a possible choice of $U(\cdot, t)$ for $t \in \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right\}$ is given above.

stochastic differential equation

$$Z_t^\varepsilon(x) = x - \int_0^t \nabla U\left(Z_s^\varepsilon, \frac{s}{2T}\right) \mathrm{d}s + \varepsilon \int_0^t g(Z_s^\varepsilon) \diamond \mathrm{d}L_s, \quad t \geqslant 0, \tag{1.4}$$

which uses the canonical Marcus integral marked by the symbol $\diamond$. For simplicity we assume that through the geometry of the potential and the definition of the vague limit of the jump measure of L a transition from the left well to the right is most likely at $(2k+1)T$, $k \in \mathbb{N}_0$ and a transition back is favoured at $2kT$, $k \in \mathbb{N}$.

Before we perform a thorough analysis of these differential equations, the first section of Chapter 3 is devoted to a model reduction describing the effective dynamics of the equations (1.3) and (1.4). Approximating two-state Markov chains with $2T$-periodic transition rates of order ε^α serve as simplification. Because of a failure to compute stochastic resonance through the use of traditional but not robust quality measures of tuning arising from spectral or information theory an approach by Herrmann and Imkeller ([19]) was chosen. Suppose the Markov chain starts at the state representing the bigger well at time $t = 0$ and the smaller one after half a period T has passed by. Then the probabilistic quantity to jump at a time point within a small interval surrounding T where transition is most likely is maximized either in ε or in T. Double stochastic resonance is documented because of a successful maximization with respect to ε respectively T although the results are not as impressive as in the case where Brownian motion serves as a perturbation term what can has its cause in the same polynomial order of the critical time scales. As the theory suggests, $2T$ must be of order $\varepsilon^{-\alpha}$. Similar quantities of the solutions of (1.3) and (1.4) are estimated in the remaining part of Chapter 3. Due to the fast growth of ∇U at infinity and the existence of a separatrix dividing the two regions of attraction of the minima a focus on bounded and reduced domains

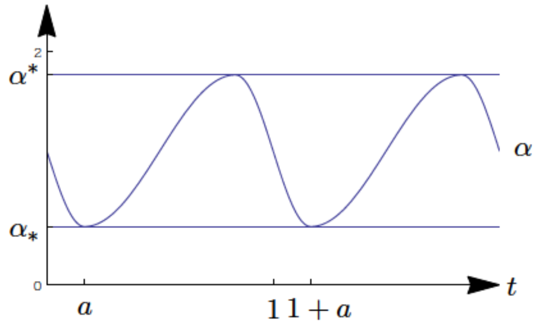


Figure 1.8: The periodically changing stability index α varies between α_* and α^*.

(excluding values near the boundary of the domains and far away from zero) is necessary. For that transitions from one bounded and reduced domain of attraction of a minimum into the other one performed by the scaled processes X^ε_{2Tt} respectively Z^ε_{2Tt} are examined where $2T$ is of the proposed order $\varepsilon^{-\alpha}$. The probability to observe such a transition at a time point within a small interval surrounding $\frac{1}{2}$ is under estimation and a similar expression as for the approximating Markov chains is obtained if localization errors are ignored and special hypotheses concerning the mass of the Lévy measure near the separatrix and on other critical sets are tolerated.

Until now a weak periodic variation of an otherwise time-independent potential always represents the periodic modulation that is necessary for stochastic resonance. In Chapter 4 the potential is kept fixed in time and a deterministic parameter of the perturbation term is chosen to be $2T$-periodic. The considered noise term belongs to the set of additive processes which is a superset of Lévy processes. Additive processes start in zero, have independent increments, are stochastically continuous, have càdlàg paths and in general do not show stationary increments like Lévy processes. Simple examples for additive processes are time changed Lévy processes. Assume $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous and increasing with $\lambda(0) = 0$ and $(L_t)_{t\geqslant 0}$ denotes a Lévy process. Then $(L_{\lambda(t)})_{t\geqslant 0}$ denotes an additive process (Example 14.5 in [44]).

Assume $\alpha\colon \mathbb{R}_+ \to [\alpha_*, \alpha^*]$ with $0 < \alpha_* < \alpha^* < 2$ is a smooth 1-periodic function with unique minimum at $a \in (0,1)$ and a unique maximum somewhere in $[0,1]$ as seen for example in Figure 1.8 and $A^T = (A^T_t)_{t\geqslant 0}$ denotes an additive process with $\alpha\left(\frac{t}{2T}\right)$-stable local characteristics for all $t \geqslant 0$. Thus the jump behaviour of A^T alters in time and represents the $2T$-periodic modulation. In Chapter 4 we capture many interesting features of the solution of the equation

$$Y^\varepsilon_t(y) = y - \int_0^t \nabla V(Y^\varepsilon_s)\,\mathrm{d}s + \varepsilon A^T_t, \tag{1.5}$$

where $V\colon \mathbb{R}^d \to \mathbb{R}_+$ is a double-well potential. Since the jump affinity of A^T is biggest for an at most small value of the stability index α, the minimal value α_* of the function α plays a significant role in finding the right time scale on which the ε-dependence of the solution is

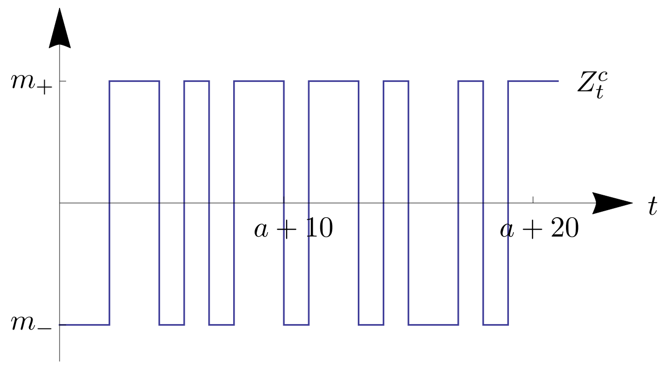


Figure 1.9: An example of a path of the limit Markov chain Z^c that attains the well minima as values and is able to leave one state when α admits its minimal value α_*.

erased and Y^ε mainly behaves like a two-state Markov chain hopping between the minima with ε-independent transition rates. Intuitively a scale of order $\varepsilon^{-\alpha_*}$ might be too small to open a time window for an observation of a transition however all scales ε^{-b} with $b > \alpha_*$ are too big and immediately cause chaotic behaviour if the stability index falls below b. The conjecture is that a scale only slightly above $\varepsilon^{-\alpha_*}$ will appear to be correct.
As in Chapter 3 at first an approximating Markov chain documents the essential behaviour of Y^ε. The probabilistic measure of a jump within a small interval surrounding $2Ta$ which is the first time the smallest stability index is attained reveals the critical time scale $T(\varepsilon) \approx \varepsilon^{-\alpha_*}\sqrt{|\log\varepsilon|}$. An analysis of the exit times of the jump diffusion guarantees the same important scale for the jump diffusion solving (1.5). We prove that $Y^\varepsilon_{t\varepsilon^{-\alpha_*}|\log\varepsilon|^{1/2}}$ in the small noise limit mainly behaves like a two-state Markov chain Z^c only allowed to jump at those time points where the stability index attains its minimal value.

Chapter 2

Mathematical framework

The introduction already gave a general overview of the topic of this work. In this chapter the underlying mathematical models for equations (1.3)-(1.5) are rigorously defined and important results from literature are presented. At first, differential equations without stochastic influence are considered, later the stochastic perturbation terms are introduced. Afterwards the uniqueness and existence of solutions of the stochastic differential equations are proven and the Markov property is discussed. Eventually Laplace's method for evaluation of integrals depending on parameters is presented due to its importance in Chapter 4.

2.1 Ordinary differential equations and periodic solutions

The stochastic differential equations (1.3), (1.4) and (1.5) arise as small amplitude perturbations of ordinary differential equations. Thus a brief analysis of the underlying gradient systems should precede the examination of the stochastic dynamics. The double-well potential V in (1.5) satisfies the conditions below:

(V1) $V \in C^2(\mathbb{R}^d, \mathbb{R}_+)$.

(V2) Stationary points of ∇V are m_-, m_+ and m_0. The eigenvalues of $\left(-\frac{\partial^2}{\partial y_i \partial y_j} V(m_-)\right)_{i,j=1}^{d}$ and $\left(-\frac{\partial^2}{\partial y_i \partial y_j} V(m_+)\right)_{i,j=1}^{d}$ are negative. Thus V has local minima at m_- and m_+. At m_0 there is a saddle point, while the Hessian matrix of V at m_0 has non zero eigenvalues.

(V3) There are $c_V^* > 0$ and $R_V^* > 0$ such that for $O_R^V \coloneqq \left\{y \in \mathbb{R}^d : V(y) \leqslant R\right\}$, $R \geqslant R_V^*$ we have $\{m_-, m_0, m_+\} \subseteq O_{R_V^*}^V \backslash \partial O_{R_V^*}^V$ and $\langle \nabla V(y), y\rangle \geqslant c_V^* \|y\|^{2+c_V^*}$ holds for all $y \in \mathbb{R}^d \backslash O_{R_V^*}^V$.

(V4) The derivatives up to order two of $\log(1 + V(y))$ are bounded.

Let $y(t) = y(t; y_0)$ be the solution of the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} y(t) = -\nabla V\left(y(t)\right), \, t \geqslant 0, \quad y(0) = y_0 \tag{2.1}$$

with initial value $y_0 \in \mathbb{R}^d$. In the following an assertion $\mathfrak{A}$ that depends on the minimum position m_- or m_+ holds simultaneously for both positions if we write $\mathfrak{A}(m_\pm)$ is true.
We shall also work with the time-dependent periodic potential U. Let the succeeding assumptions on the geometry of U be fulfilled.

(U1) $U \in C^{2,1}(\mathbb{R}^d \times \mathbb{R}_+, \mathbb{R}_+)$ and $U(x,t) = U(x,t+1)$ for all $x \in \mathbb{R}^d$, $t \geqslant 0$.

(U2) It holds $\nabla U(x,t) = 0$ if and only if $x \in \{m_-(t), m_+(t), m_0(t)\}$. The eigenvalues of $\left(-\frac{\partial^2}{\partial x_i \partial x_j} U(m_\pm(t), t)\right)_{i,j=1}^d$ are negative and bounded away from zero uniformly in t. The corresponding Hessian matrix of the saddle position $m_0(t)$ is indefinite, while again uniform boundedness from zero is valid for the eigenvalues.

(U3) There are $c_U^* > 0$ and $R_U^* > 0$ such that for $O_R^U := \left\{x \in \mathbb{R}^d : \|x\|^{2+c_U^*} \leqslant R\right\}$, $R \geqslant R_U^*$ we have $\{m_i(t) : i \in \{-,0,+\}, t \geqslant 0\} \subseteq O_{R_U^*}^U \backslash \partial O_{R_U^*}^U$ and $\langle \nabla U(x,t), x\rangle \geqslant c_U^* \|x\|^{2+c_U^*}$ holds for all $x \in \mathbb{R}^d \backslash O_{R_U^*}^U$ and $t \geqslant 0$.

(U4) The derivatives $\frac{\partial}{\partial t}$, $\frac{\partial}{\partial x_i}$ and $\frac{\partial^2}{\partial x_i \partial x_j}$ for $i,j \in \{1,\dots,d\}$ of $\log(1+U(x,t))$ are uniformly bounded in $t \geqslant 0$.

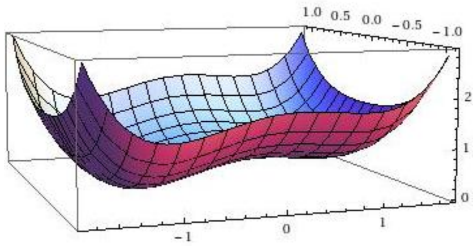
A possible 1-dimensional example is $U(x,t) = \frac{x^4}{4} - \frac{x^2}{2} + c\,x \cos 2\pi t + C$ for $x \in \mathbb{R}$, $t \geqslant 0$ and fixed $c, C > 0$. Its natural extension to d dimensions can look like:

$$U(x_1,\dots,x_d,t) = \frac{x_1^4}{4} - \frac{x_1^2}{2} + c\,x_1 \cos 2\pi t + \sum_{i=2}^d \frac{x_i^2}{2}(1+x_i^2) + C \tag{2.2}$$

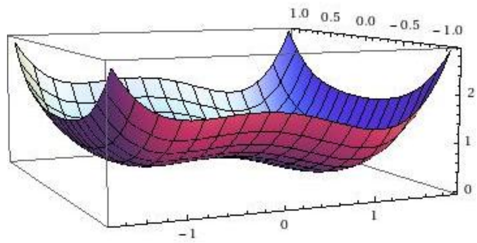
for $(x_1,\dots,x_d) \in \mathbb{R}^d$ and $t \geqslant 0$. The potential U is symmetric for $t = (2k-1)\frac{1}{4}$, $k \in \mathbb{N}$, and the well sizes are most different at $t = \frac{k}{2}$, $k \in \mathbb{N}_0$. Consider the solution $x(t; x_0, t_0)$, $t \geqslant t_0$, of the time non-autonomous ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} x(t) = -\nabla U\left(x(t), \frac{t}{2T}\right), \; t \geqslant t_0, \quad x(t_0) = x_0, \tag{2.3}$$

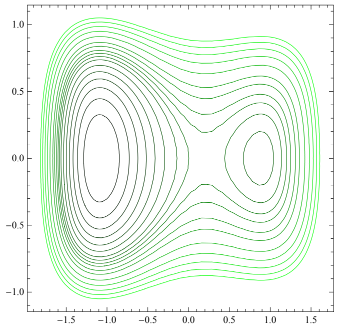
with initial value $x_0 \in \mathbb{R}^d$ and starting time $t_0 \geqslant 0$.
Continuity and the local Lipschitz property in the space variable of ∇U and ∇V suffice to prove existence and uniqueness of solutions x and y (Theorem 3.1 in Chapter 1 of [18]). The steepness of the potential makes a divergence of the solutions impossible and we can work with $x(t)$ and $y(t)$ for all $t \geqslant 0$.
In Chapter 3 and 4 we mainly concentrate on stochastic differential equations with starting values from bounded sets. It is natural to work with invariant sets here. For the gradient system (2.1) the union O_R^V of all level sets below a certain level R is invariant (Theorem 16.9 in [1]). The non-autonomous equation (2.3) also has invariant sets O_R^U due to assumption (U3) and Satz 7.2.1 in [45]. Additionally the conditions (V3) and (U3) guarantee a uniform upper bound for the return time of the deterministic solutions to $O_{R^*}^V$ respectively $O_{R^*}^U$.



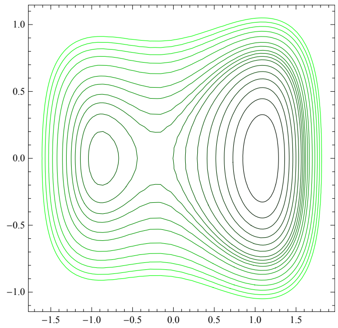
(a) $t = 0$



(b) $t = \frac{1}{2}$



(c) Level sets of $U(x, 0)$



(d) Level sets of $U(x, \frac{1}{2})$

Figure 2.1: The potential U given in formula (2.2) for $d = 2$, $c = 0.2$, $C = 0.5$ and $t = 0$ and $t = \frac{1}{2}$. The dark level sets reflect the greater depth of the well.

While the minimum positions m_- and m_+ are exponentially stable, the inhomogeneity of (2.3) complicates the analysis of the convergence of its solution as $t \to \infty$. If $2T$ is chosen very big, then U varies very slowly while the attraction of the region around the well bottoms is quite strong. Thus for nearly all starting points the solution x of (2.3) very fast enters the neighbourhood of one minimum and always follows it with a small $\frac{1}{2T}$-dependent delay. In [6] and [4] Berglund and Gentz intensively examined slow-fast systems of the form

$$\frac{1}{2T}\frac{\mathrm{d}}{\mathrm{d}s}\hat{x}(s) = -\nabla U(\hat{x}(s), s),\ s \geqslant \frac{t_0}{2T} =: s_0, \quad \hat{x}\left(\frac{t_0}{2T}\right) = x_0.$$

Because of the periodicity of U in time the question of existence of periodic solutions arises.

Lemma 2.1. ([6] Theorem 2.18.) *There exist invariant solutions $\hat{x}_i^{inv}$ for $i \in \{-, 0, +\}$ with $\hat{x}_i^{inv}(s) = m_i(s) + \mathrm{O}(\frac{1}{2T})$.*

From $s = \frac{t}{2T}$ one can derive $x\left(t; \hat{x}_i^{inv}\left(\frac{t_0}{2T}\right), \frac{t_0}{2T}\right) = \hat{x}_i^{inv}\left(\frac{t}{2T}\right)$ for all $t \geqslant 0$.

Definition 2.2. *Define the $2T$-periodic function $p_i(t) := \hat{x}_i^{inv}\left(\frac{t}{2T}\right)$ for $i \in \{-, 0, +\}$.*

In Section 2.4 we will choose $2T$ of order $\varepsilon^{-\alpha}$ according to a special parameter α of the perturbation process L in formulae (1.3) and (1.4). Hence the norm $\left\|m_\pm(s) - p_\pm\left(\frac{s}{2T}\right)\right\|$ will be of order ε^α.

2.2 Stochastic processes

The most important ingredients of the perturbation terms used in (1.3)-(1.5) are Lévy processes and additive processes which are defined in the following.

Definition 2.3. *The stochastic process $(L_t)_{t\geqslant 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with $L_0 = 0$ a.s. denotes a Lévy process on $\mathbb{R}^d$ if the following conditions hold.*

(i) For $0 \leqslant t_1 < \ldots < t_n$, $n \in \mathbb{N}$, the increments $L_{t_n} - L_{t_{n-1}}, \ldots, L_{t_1} - L_{t_0}$ are independent.

(ii) The law of $L_{t+s} - L_s$ is independent of s for all $s, t \geqslant 0$ and $L_{t+s} - L_s \stackrel{\mathrm{d}}{=} L_t$.

(iii) L is stochastically continuous, that is $\lim_{s\to t} \mathbb{P}(\|L_t - L_s\| > \varepsilon) = 0$ for all $\varepsilon, t \geqslant 0$.

(iv) The paths of L are a.s. càdlàg. Thus there exists $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ and $L_t(\omega)$ is right-continuous and has left limits in t for $\omega \in \Omega_0$.

A process $A = (A_t)_{t\geqslant 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with $A_0 = 0$ a.s. that satisfies only (i),(iii) and (iv) is called additive process.

Definition 2.4. *A measure ν on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is called Lévy measure if the following is fulfilled:*

$$\nu(\{0\}) = 0, \qquad \int_{\mathbb{R}^d} \left(\|x\|^2 \wedge 1\right) \mathrm{d}\nu(x) < \infty.$$

Additive processes are strongly linked to infinitely divisible distributions.

Theorem 2.5. ([42] Theorem 9.8.) *Assume $(A_t)_{t\geqslant 0}$ is an additive process. Then there exist Lévy measures μ_t, $t \geqslant 0$, vectors $a_t \in \mathbb{R}^d$, $t \geqslant 0$, and positive-semidefinite $d \times d$ matrices Σ_t, $t \geqslant 0$ such that*

$$\mathbb{E}\, \mathrm{e}^{\mathrm{i}\langle x, A_t\rangle} = \exp\left(\mathrm{i}\,\langle x, a_t\rangle - \frac{1}{2}\langle x, \Sigma_t x\rangle + \int_{\mathbb{R}^d} \left(\mathrm{e}^{\mathrm{i}\langle x,y\rangle} - 1 - \mathrm{i}\,\langle x, y\rangle\, \mathbf{1}_{B_1(0)}(y)\right) \mu_t(\mathrm{d}y)\right)$$

is true for all $x \in \mathbb{R}^d$ and $t \geqslant 0$ while the following conditions are valid:

(i) it holds $a_0 = 0$, $\Sigma_0 = 0$ and $\mu_0 = 0$;

(ii) for $0 \leqslant s \leqslant t$ we have $\langle x, \Sigma_s x\rangle \leqslant \langle x, \Sigma_t x\rangle$ and $\mu_s(B) \leqslant \mu_t(B)$ for all $x \in \mathbb{R}^d$, $B \in \mathcal{B}(\mathbb{R}^d)$;

(iii)

$$a_s \to a_t, \qquad \langle x, \Sigma_s x\rangle \to \langle x, \Sigma_t x\rangle, \qquad \mu_s(B) \to \mu_t(B),$$

for Borel sets $B \subseteq \left\{x \in \mathbb{R}^d : \|x\| > \delta\right\}$ for some $\delta > 0$, as $|s - t| \to 0$, $t \geqslant 0$.

Conversely, for a family of triplets $(a_t, \Sigma_t, \mu_t)_{t\geqslant 0}$ fulfilling (i)-(iii) there exists an additive process A with characteristic function given above. It is unique up to identity in law.

The family $(a_t, \Sigma_t, \mu_t)_{t\geqslant 0}$ is called system of generating triplets. In the case of Lévy processes it equals to $(ta, t\Sigma, t\nu)_{t\geqslant 0}$ (short: (a, Σ, ν)) with $a \in \mathbb{R}^d$, a positive-semidefinite $d \times d$ matrix Σ, and a Lévy measure ν.

Remark 2.6. ([42] Remark 9.9.) The Lévy measures $(\mu_t)_{t\geqslant 0}$ uniquely determine a measure m on $\mathbb{R}^d \times [0,\infty)$ with

$$m(B \times [0,t]) = \mu_t(B)$$

for $t \geqslant 0$ and $B \in \mathcal{B}(\mathbb{R}^d)$ and m satisfies $m(\mathbb{R}^d \times \{t\}) = 0$ and

$$\int_{\mathbb{R}^d \times [0,t]} \left(\|x\|^2 \wedge 1\right) m(\mathrm{d}(x,s)) < \infty$$

for all $t \geqslant 0$. Since $m\left((\mathbb{R}^d \times \{0\}) \cup (\{0\} \times \mathbb{R}_+)\right) = 0$, it is a measure on $\mathbb{R}^d \setminus \{0\} \times (0,\infty)$, too. If we consider a Lévy process, m above is the product measure of the Lévy measure and the Lebesgue measure on $\mathbb{R}_+$.

Example 2.7. ([44] pp. 458-459) *Assume $\gamma\colon \mathbb{R}_+ \to \mathbb{R}^d$ is a function with finite variation on compacts, σ is a matrix-valued function on $\mathbb{R}_+$ while $\sigma(t)$ is symmetric and $\int_0^t \sigma_{ij}^2(s)\,\mathrm{d}s$ is finite for all $i,j = 1,\ldots,d$, $t \geqslant 0$ and ν_t is a Lévy measure for all $t \geqslant 0$, then (a_t, Σ_t, μ_t) with a_t, Σ_t and μ_t given by*

$$\begin{aligned} a_t &= \int_0^t \gamma(s)\,\mathrm{d}s, \\ \Sigma_t &= \int_0^t \sigma^2(s)\,\mathrm{d}s, \\ \mu_t(B) &= \int_0^t \nu_s(B)\,\mathrm{d}s, \end{aligned}$$

is an admissible generating triplet for an additive process. The triplet $(\gamma(t), \sigma(t), \nu_t)$ is called local characteristic.

Remark 2.8. In Chapter 4 we will work with an additive process given according to the last example with ν_t defined by

$$\nu_t(\mathrm{d}x) = \frac{c(t)}{\|x\|^{d+\alpha(t)}}\,\mathrm{d}x, \quad t \geqslant 0,$$

for smooth and periodic functions c and α.

Multiplicative noise terms in (1.3) and (1.4) and computations of integrals with respect to additive processes in Chapter 4 require that all possible integrators belong to the set of semimartingales.

Definition 2.9. *A semimartingale Y on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is an adapted, càdlàg process that admits the decomposition $Y = Y_0 + M + V$, where Y_0 is finite-valued and $\mathcal{F}_0$-measurable, M is a local martingale with $M_0 = 0$, and V is a process of finite variation on compacts with $V_0 = 0$.*

Lemma 2.10. ([27] Corollary 5.11 in Chapter 2) *An additive process A with triplet $(a_t, \Sigma_t, \mu_t)_{t\geqslant 0}$ is a semimartingale if and only if a_t is of bounded variation on finite intervals.*

Well-known processes like the Brownian motion, the Poisson process or in general all Lévy processes belong to the nice class of strong Markov processes. The independence of the future and the past of a process given the present state is called Markov property. If stopping times are allowed as time points, the word "strong" is added.

Definition 2.11. *The $\mathbb{R}^d$-valued, $\mathbb{F}$-adapted process X is a Markov process with respect to $\mathbb{F}$ if for all measurable and bounded functions f and $t, s \geqslant 0$*

$$\mathbb{E}\left(f(X_{t+s})|\mathcal{F}_t\right) = \mathbb{E}\left(f(X_{t+s})|X_t\right) \ \ a.s.$$

If additionally for all $B \in \mathcal{B}(\mathbb{R}^d)$ the expression $\mathbb{P}(X_{t+s} \in B|X_t)$ is independent of s, X is a time-homogeneous Markov process. A homogeneous Markov process X is called strong Markov if for all bounded stopping times τ the succeeding equation is true

$$\mathbb{E}\left(f(X_{\tau+s})|\mathcal{F}_\tau\right) = \mathbb{E}\left(f(X_{\tau+s})|X_\tau\right) \ \ a.s.$$

Lemma 2.12. ([2] Theorem 2.2.11.), ([42] Theorem 10.4.) *Lévy processes are strong Markov processes with respect to their completed natural filtration. Additive processes are Markov processes with respect to their completed natural filtration.*

The next highlight is the Lévy-Itô decomposition of additive processes which requires further definitions.

Definition 2.13. *Assume $(E, \mathcal{E}, m)$ is a σ-finite measure space. The family of random variables $(N(B))_{B\in\mathcal{E}}$ is a Poisson random measure with intensity measure m if*

(i) $N(B) \in \mathbb{N}_0 \cup \{\infty\}$,

(ii) $N(B)$ *is Poisson distributed with parameter* $m(B)$,

(iii) if $B_1, \ldots, B_n \in \mathcal{E}$ are disjoint for some $n \in \mathbb{N}$, then $N(B_1), \ldots, N(B_n)$ are independent,

(iv) $N(\cdot, \omega)$ *is a measure on* $(E, \mathcal{E})$.

Lemma 2.14. ([42] Proposition 19.5.) *Let $(E, \mathcal{E}, m)$ be a σ-finite measure space with $m(E) < \infty$, $(N(B))_{B\in\mathcal{E}}$ denotes a Poisson random measure with intensity m and $f\colon E \to \mathbb{R}^d$ is measurable. Then*

$$X(\omega) := \int_E f(x)N(\mathrm{d}x, \omega),$$

is compound Poisson distributed and $\int_E \|f(x)\|^2\, m(\mathrm{d}x) < \infty$ implies $\mathbb{E}\|X\|^2 < \infty$ and

$$\mathbb{E}X = \int_E f(x)\, m(\mathrm{d}x),$$

$$\mathbb{E}\|X - \mathbb{E}X\|^2 = \int_E \|f(x)\|^2\, m(\mathrm{d}x).$$

If $B_1, \ldots, B_n \in \mathcal{E}$, $n \in \mathbb{N}$, are disjoint, then $\int_{B_i} f(x)N(\mathrm{d}x, \omega)$ for $i = 1, \ldots, n$ are independent.

Remark 2.15. The jump times of an additive process A with triplet $(a_t, \Sigma_t, \mu_t)_{t \geqslant 0}$ and their sizes can be described through a Poisson random measure N^A which plays an important role in the Lévy-Itô decomposition. Assume $A_{\cdot}(\omega)$ is càdlàg for $\omega \in \Omega_0$. Define

$$N^A(B, \omega) = \begin{cases} \# \{s \geqslant 0 : (A_s(\omega) - A_{s-}(\omega), s) \in B\}, & \omega \in \Omega_0, \\ 0, & \omega \in \Omega_0^c, \end{cases}$$

for $B \in \mathcal{B}(\mathbb{R}^d \setminus \{0\} \times (0, \infty))$. The measure m on $\mathcal{B}(\mathbb{R}^d \setminus \{0\} \times (0, \infty))$ defined as in Remark 2.6 denotes the intensity of N^A.

Theorem 2.16. ([42] Theorem 19.2.) *Let A be an additive process with a generating triplet $(a_t, \Sigma_t, \mu_t)_{t \geqslant 0}$. Then $(N^A(B))_{B \in \mathcal{B}(\mathbb{R}^d \setminus \{0\} \times (0,\infty))}$ is a Poisson random measure on $\mathbb{R}^d \setminus \{0\} \times (0, \infty)$ with intensity measure m defined in Remark 2.6. There exist $\Omega_i \in \mathcal{F}$ with $\mathbb{P}(\Omega_i) = 1$ for $i = 1, 2$ such that the following can be defined for $t \geqslant 0$*

(i)

$$\begin{aligned} A_t^1(\omega) = \lim_{\varepsilon \to 0} & \int_{\{\|x\| \in (\varepsilon, 1]\} \times (0,t]} x \left(N^A(\mathrm{d}(x,s), \omega) - m(\mathrm{d}(x,s)) \right) \\ & + \int_{\{\|x\| > 1\} \times (0,t]} x \, N^A(\mathrm{d}(x,s), \omega), \quad \omega \in \Omega_1, \end{aligned}$$

(ii)

$$A_t^2(\omega) = A_t(\omega) - A_t^1(\omega), \quad \omega \in \Omega_2.$$

The convergence in (i) is uniform in t for all bounded intervals. The processes A^1 and A^2 are independent and additive with generating triplet $(0, 0, \mu_t)$ respectively $(a_t, \Sigma_t, 0)$.

The process A^1 is often referred to as jump part, A^2 is called continuous part. Due to the product structure of the space $\mathbb{R}^d \setminus \{0\} \times (0, \infty)$ we often write $(\mathrm{d}x, \mathrm{d}s)$ instead of $\mathrm{d}(x, s)$. In the stochastic differential equation (1.5) an additive process $A = A^T$ serves as noise term and $N^A = N^{A^T}$ is the associated Poisson random measure (Remark 2.15) with intensity m. In Chapter 4 also integrals of random integrands with respect to the compensated Poisson random measure $\widetilde{N}^A := N^A - m$ will be considered. Thus we will have to define stochastic integrals of the form

$$\int_0^t \int_{\|x\| \in (0,1)} f(x,s) \left(N^A(\mathrm{d}x, \mathrm{d}s) - m(\mathrm{d}(x,s)) \right)$$

for a function $f \colon \Omega \times \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ following Applebaum ([2] Section 4.2).

Definition 2.17. *The function $f \colon \Omega \times \mathbb{R}^d \times [0, t] \to \mathbb{R}$ is called simple if*

$$f(x,t) = \sum_{i,j=1}^{n,m} f_{ij} \, \mathbf{1}_{(t_i, t_{i+1}]}(t) \, \mathbf{1}_{B_j}(x)$$

holds where $0 \leqslant t_1 \leqslant \ldots \leqslant t_n \leqslant t$, f_{ij} is a $\mathcal{F}_{t_i}$-measurable, bounded random variable, and Borel sets $B_1, \ldots, B_m \in B_1(0)$ are disjoint with $m(B_j \times [0,t]) < \infty$.

The first approach applies L_2-theory using the definition

$$I_t(f) := \int_0^t \int_{\|x\|\in(0,1)} f(x,s)\, \widetilde{N}^A(\mathrm{d}x,\mathrm{d}s) := \sum_{i,j=1}^{n,m} f_{ij}\, \widetilde{N}^A(B_j \times (t_i \wedge t, t_{i+1} \wedge t]),$$

for a simple function f as above. The process $I_t(f)$ is a square-integrable martingale and fulfills the isometry formula (Lemma 4.2.2 in [2])

$$\mathbb{E}|I_t(f)|^2 = \mathbb{E} \int_0^t \int_{\|x\|\in(0,1)} f(x,s)^2\, m(\mathrm{d}(x,s)).$$

Together with the fact that simple functions are dense in the set of all predictable square-integrable functions this enables us to formulate the next definition.

Definition 2.18. *Let f be predictable with*

$$\mathbb{E} \int_0^t \int_{\|x\|\in(0,1)} f(x,s)^2\, m(\mathrm{d}(x,s)) < \infty$$

and f^n are simple functions with

$$\lim_{n\to\infty} \mathbb{E} \int_0^t \int_{\|x\|\in(0,1)} |f^n(x,s) - f(x,s)|^2\, m(\mathrm{d}(x,s)) = 0.$$

Then define

$$I_t(f) := \int_0^t \int_{\|x\|\in(0,1)} f(x,s)\, \widetilde{N}^A(\mathrm{d}x,\mathrm{d}s) := L_2 \text{ - }\lim_{n\to\infty} \int_0^t \int_{\|x\|\in(0,1)} f^n(x,s)\, \widetilde{N}^A(\mathrm{d}x,\mathrm{d}s).$$

Lemma 2.19. ([2] Theorem 4.2.3.) *The stochastic process $(I_t(f))_{t\geqslant 0}$ defined above is a square-integrable martingale and satisfies*

$$\mathbb{E}\|I_t(f)\|^2 = \int_0^t \int_{\|x\|\in(0,1)} \mathbb{E}\|f(x,s)\|^2\, m(\mathrm{d}(x,s)).$$

Remark 2.20. ([2] pp. 225-229, 230-231) Simple functions are also dense in the set of predictable functions with $\mathbb{P}(\int_0^t \int_{\|x\|\in(0,1)} |f(x,s)|^2\, m(\mathrm{d}(x,s)) < \infty) = 1$ which makes an extension of $I_t(\cdot)$ to those functions possible. The new definition yields a local martingale. It is even possible to generalize the integral term to integrands f with $\mathbb{P}(\int_0^t \int_{\|x\|\in(0,1)} |f(x,s)|\, m(\mathrm{d}(x,s)) < \infty) = 1$ which then creates a local martingale that is even a martingale if $\mathbb{E} \int_0^t \int_{\|x\|\in(0,1)} |f(x,s)|\, m(\mathrm{d}(x,s)) < \infty$.

Lemma 2.21. *Let N^A be a Poisson random measure on $\mathbb{R}^d \setminus \{0\} \times (0,\infty)$ corresponding to the jumps of an additive process A with compensated version $\widetilde{N}^A$. Define*

$$X_t^i = \int_0^t \int_{\|x\|\in(0,1)} g^i(x,s)\, \widetilde{N}^A(\mathrm{d}x,\mathrm{d}s), \quad i = 1,2$$

for predictable and square-integrable g^i. Then the quadratic covariation process is

$$[X^1, X^2]_t = \int_0^t \int_{\|x\|\in(0,1)} g^1(x,s) g^2(x,s)\, N^A(\mathrm{d}x,\mathrm{d}s).$$

Proof. The decomposition of the quadratic covariation into a continuous and a purely discontinuous part and the knowledge of the jump behaviour of X^i proves the statement. □

2.3 The heavy tail property

In Chapter 3 not arbitrary Lévy processes serve as integrators in the noise terms. To obtain heavy tails in comparison to the exponentially light Gaussian tails we assume the Lévy measure is of regular variation which is specified in short in this section. Recommended literature on this topic are works of Hult and Lindskog ([21],[22]), Resnick ([40]) and Lindskog ([33]).

Definition 2.22. *A measurable function $l\colon \mathbb{R}_+ \to \mathbb{R}_+$ is called slowly varying (at ∞) if*

$$\lim_{t\to\infty} \frac{l(\lambda t)}{l(t)} = 1, \text{ for all } \lambda > 0.$$

A measurable function $r\colon \mathbb{R}_+ \to \mathbb{R}_+$ with

$$\lim_{t\to\infty} \frac{r(\lambda t)}{r(t)} = \lambda^{\alpha}, \text{ for all } \lambda > 0$$

and $\alpha \in \mathbb{R}$ is called regularly varying (at ∞) with index α.

Theorem 2.23. ([7] Theorem 1.4.1.) *Every regularly varying function r with index α admits a representation $r(x) = x^{\alpha} l(x)$, $x \in \mathbb{R}_+$, with a slowly varying function l.*

Example 2.24. *Prominent examples of distributions on $\mathbb{R}$ with heavy tails are α-stable distributions which fulfill $\lim_{x\to\infty} x^{\alpha}\mathbb{P}(X > x) = c_{\alpha} \in (0,\infty)$ for $\alpha \in (0,2)$ (cf. [2] Section 1.5.4). A Lévy process $X = (X_t)_{t\geqslant 0}$ being α-stable at time $t = 1$ is called α-stable process. The corresponding Lévy measure ν of a symmetric α-stable Lévy process satisfies $\mathrm{d}\nu(x) = c\,|x|^{-1-\alpha}\mathrm{d}x$, for $x \neq 0$ and $c > 0$ (Remark 14.4 in [42]) and hence has infinite intensity $\nu(\mathbb{R}) = +\infty$. It holds $\frac{\nu(\{x\in\mathbb{R}:|x|\geqslant\lambda t\})}{\nu(\{x\in\mathbb{R}:|x|\geqslant t\})} = \lambda^{-\alpha}$ for $t, \lambda > 0$. Thus the tails $\nu(\{x \in \mathbb{R} : |x| \geqslant \cdot\})$ are regularly varying with index $-\alpha$.*

The definition of regularly varying measures requires the familiarity with Radon measures.

Definition 2.25. *A measure μ on $\mathcal{B}(\bar{\mathbb{R}}^d\setminus\{0\})$ is called Radon measure if $\mu(B) < \infty$ for all relative compact sets $B \in \mathcal{B}(\bar{\mathbb{R}}^d\setminus\{0\})$.*

Definition 2.26. *A Radon measure ν is called regularly varying with index $-\alpha < 0$ if and only if there exists a non-zero Radon measure μ such that for all $\lambda > 0$*

$$\lim_{t\to\infty} \frac{\nu(\lambda t B)}{\nu\left(t\left\{x \in \mathbb{R}^d : \|x\| \geqslant 1\right\}\right)} = \lambda^{-\alpha} \lim_{t\to\infty} \frac{\nu(tB)}{\nu\left(t\left\{x \in \mathbb{R}^d : \|x\| \geqslant 1\right\}\right)} = \lambda^{-\alpha}\mu(B)$$

for all $B \in \mathcal{B}(\bar{\mathbb{R}}^d)$ with $\mu(\partial B) = 0$ and $0 \notin \bar{B}$. The reference set $\left\{x \in \mathbb{R}^d : \|x\| \geqslant 1\right\}$ could be replaced by any Borel set which is bounded away from zero.

Remark 2.27. The definition of regular variation of Radon measures is usually formulated through vague convergence of measures but Portmanteau theorem (Theorem 3.1 in [14]) immediately links this convergence to the convergence of measures of relative compact sets. Thus we preferred this as definition.

Remark 2.28. An immediate consequence of Definition 2.26 is $\mu(\lambda B) = \lambda^{-\alpha}\mu(B)$ for all $\lambda > 0$. This scaling property implies the lack of atoms, $\mu(rS^{d-1}) = 0$ for all spheres rS^{d-1} centered at 0 with radius $r > 0$ and many other nice properties listed in Theorem 1.8 in Chapter 1 of [33]. Regular variation of a Radon measure ν with index $-\alpha$ entails regular variation of $\nu\left(\{x \in \mathbb{R}^d : \|x\| \geqslant \cdot\}\right)$ and thus $\nu\left(\{x \in \mathbb{R}^d : \|x\| \geqslant \varepsilon^{-1}\}\right)$ decreases like $\varepsilon^{\alpha} l(\varepsilon^{-1})$ for a slowly varying l as ε tends to zero.

2.4 Stochastic differential equations

All necessary ordinary differential equations were analysed in Section 2.1 and now perturbation terms are added to obtain stochastic differential equations. At first definitions of the used noise terms follow and afterwards existence and uniqueness results for solutions of the created stochastic differential equations are presented.

2.4.1 Perturbation terms and stochastic integration

We formulate the following assumptions concerning the noise terms in equation (1.3) and (1.4) and the period length $2T$.

(N1) Let $L = (L_t)_{t\geqslant 0}$ denote a d-dimensional Lévy process with characteristic triplet (a, Σ, ν) and ν is regularly varying with index $-\alpha < 0$ and a Radon measure μ as limit measure according to Definition 2.26.

(N2) The function g on $\mathbb{R}^d$ is matrix-valued and every component g_{ij} is continuous and bounded and has bounded and Lipschitz continuous derivatives.

(T) The period length $2T$ is coupled to the noise intensity ε as: $2T = \frac{c_{per}}{\nu(B^c_{\varepsilon^{-1}}(0))} = \frac{c_{per}}{\varepsilon^{\alpha} l(\varepsilon^{-1})}$ for some $c_{per} > 0$ and l slowly varying.

Remark 2.29. After the main results in Chapter 3 (Theorem 3.38 and Theorem 3.43) the constant c_{per} will be chosen such that an at most periodic behaviour of the jump diffusion X^{ε} solving equation (1.3) respectively Z^{ε} solving equation (1.4) can be observed. The obtained optimal choice of $T(\varepsilon)$ will coincide with the result for the reduced Markov chain model with time-continuous infinitesimal generator considered in Subsection 3.1.2 (Proposition 3.9 for $k = 0$).

The perturbation term of the equation (1.3) uses an Itô integral of g with respect to a d-dimensional Lévy process. A nice reference to the topic of stochastic integration with respect to the wide class of semimartingales to which Lévy processes belong (Lemma 2.10) is the book of Protter ([39]). Many properties of the stochastic integral are presented there and related subjects like stochastic differential equations are discussed. A frequently applied consequence of the Itô calculus which is called Itô formula is given below.

Theorem 2.30. ([44] Proposition 8.19.) *Assume X denotes a d-dimensional semimartingale and the function $f\colon \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$ is twice differentiable with $(f_{x_1}, \dots, f_{x_d}, f_t) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_d}, \frac{\partial f}{\partial t}\right)$ and $f_{x_i,x_j} = \frac{\partial^2}{\partial x_i \partial x_j} f$. Then $f(X_t, t)$ is a semimartingale and it holds*

$$\begin{aligned} f(X_t,t) = f(X_0,0) &+ \int_0^t f_t(X_s,s)\,\mathrm{d}s + \frac{1}{2}\sum_{i,j=1}^{d}\int_0^t f_{x_i,x_j}(X_{s-},s)\,\mathrm{d}[X^i,X^j]_s^c \\ &+ \sum_{j=1}^{d}\int_0^t f_{x_j}(X_{s-},s)\mathrm{d}X_s^j + \sum_{0<s\leqslant t}\left(f(X_s,s) - f(X_{s-},s) - \sum_{j=1}^{d} f_{x_j}(X_{s-},s)\Delta X_s^j\right). \end{aligned}$$

In comparison to that in equation (1.4) the canonical Marcus integral is used which is only defined for integrands which solve a stochastic differential equation. Pages 272-274 of [2] serve as introduction and [32] is a comprehensive reference to this subject.

Definition 2.31. *Assume g satisfies* (N2). *Let Y denote a d-dimensional semimartingale and X is the solution of the stochastic differential equation*

$$X_t = X_0 + \int_0^t g(X_s) \diamond \mathrm{d}Y_s,$$

while X_0 is $\mathcal{F}_0$-measurable and the involved integral is understood as given in the following

$$\begin{aligned} \int_0^t g(X_s)\diamond \mathrm{d}Y_s = &\int_0^t g(X_{s-})\,\mathrm{d}Y_s + \frac{1}{2}\int_0^t g'(X_s)g(X_s)\,\mathrm{d}[Y,Y]_s^c \\ &+ \sum_{s\leqslant t}\left(\varphi(1;X_{s-},\Delta Y_s) - X_{s-} - g(X_{s-})\Delta Y_s\right), \end{aligned}$$

and $\varphi(\cdot;x,y)$ solves the ordinary differential equation

$$\frac{\mathrm{d}}{\mathrm{d}r}\varphi(r;x,y) = g(\varphi(r;x,y))y, \quad \varphi(0;x,y) = x.$$

Remark 2.32. Lipschitz continuity of g verifies the existence and uniqueness of φ. From Taylor's Theorem we derive that $\|\varphi(1;X_{s-},\Delta Y_s) - X_{s-} - g(X_{s-})\Delta Y_s\|$ is of order $\|\Delta Y_s\|^2$ and hence absolute convergence of the sum is guaranteed. Under the Lipschitz conditions above the existence and uniqueness of a solution X considered in Definition 2.31 which is càdlàg, a semimartingale, and satisfies the strong Markov property results from Theorem 3.2 and 5.1 of [32].

Remark 2.33. The perturbation term in (1.4) equals to

$$\begin{aligned} \varepsilon\int_0^t g(Z_s^\varepsilon)\diamond \mathrm{d}L_s = &\,\varepsilon\int_0^t g(Z_{s-}^\varepsilon)\,\mathrm{d}L_s^c + \frac{\varepsilon^2}{2}\int_0^t g'(Z_s^\varepsilon)g(Z_s^\varepsilon)\,\mathrm{d}[L,L]_s^c + \varepsilon\int_0^t g(Z_{s-}^\varepsilon)\,\mathrm{d}L_s^d \\ &+ \sum_{s\leqslant t}\left(\varphi(1;Z_{s-}^\varepsilon,\varepsilon\Delta L_s) - Z_{s-}^\varepsilon - \varepsilon g(Z_{s-}^\varepsilon)\Delta L_s\right), \end{aligned}$$

where the i-th component of the second integral is

$$\frac{\varepsilon^2}{2}\sum_{k,l,j=1}^{d}\int_0^t \frac{\partial}{\partial x_k} g_{ij}(Z_s^\varepsilon)g_{kl}(Z_s^\varepsilon)\,\mathrm{d}[L^l,L^j]_s^c.$$

A jump of Z^ε does not simply arise from $g(Z^\varepsilon_{s-})\Delta L_s$ as in the Itô case. Instead of that a jump of L causes Z^ε to move from Z^ε_{s-} to $\varphi(1; Z^\varepsilon_{s-}, \varepsilon\Delta L_s)$ along the vector field $\varepsilon g(\cdot)\Delta L_s$. One outstanding nice consequence of this type of stochastic integral is the usual chain rule which holds under special conditions. Other remarkable advantages are also listed in [32]. At first the integral of row vectors is introduced.

Definition 2.34. *Let $f \in C^1(\mathbb{R}^d, \mathbb{R}^k)$ and g and Y are given according to Definition 2.31. Assume $(X_t)_{t\geqslant 0}$ satisfies*

$$X_t = X_0 + \int_0^t g(X_s) \diamond \mathrm{d}Y_s.$$

Then define

$$\int_0^t f(X_s) \diamond \mathrm{d}Y_s = \int_0^t f(X_{s-})\,\mathrm{d}Y_s + \frac{1}{2}\,\mathrm{Trace}\left(\int_0^t f'(X_s)\,\mathrm{d}[Y,Y]^c_s\, f(X_s)^T\right) + \sum_{0<s\leqslant t}\left(\int_0^1 \left(f(\varphi(r; X_{s-}, \Delta Y_s)) - f(X_{s-})\right)\mathrm{d}r\right)\Delta Y_s.$$

Proposition 2.35. ([32] Proposition 4.2.) *Assume g is a continuously differentiable, matrix-valued function on $\mathbb{R}^d$ which has Lipschitz components and Lipschitz continuous derivatives. The process Y denotes a semimartingale, X_0 is $\mathcal{F}_0$-measurable and $(X_t)_{t\geqslant 0}$ solves*

$$X_t = X_0 + \int_0^t g(X_s) \diamond \mathrm{d}Y_s.$$

Let $\psi \in C^2(\mathbb{R}^d)$. Then it holds

$$\psi(X_t) = \psi(X_0) + \int_0^t \psi'(X_s) g(X_s) \diamond \mathrm{d}Y_s.$$

The perturbation process A^T occuring in (1.5) is an additive process. Assume, it has local characteristics $(\gamma(t), \sigma(t), \nu_t)$ and assume $c_A > 0$. Later c_A is chosen to be ε-dependent (Chapter 4) but this is unimportant now. The process A^T admits the decomposition

$$\begin{aligned} A^T_t &= \int_0^t \gamma(s)\,\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}W_s + \int_0^t \int_{\|x\|<c_A} x\left(N^{A^T}(\mathrm{d}x, \mathrm{d}s) - \nu_{s/2T}(\mathrm{d}x)\mathrm{d}s\right) \\ &\quad + \int_0^t \int_{\|x\|\geqslant c_A} x\, N^{A^T}(\mathrm{d}x, \mathrm{d}s) \\ &=: \widetilde{A}^T_t + \int_0^t \int_{\|x\|\geqslant c_A} x\, N^{A^T}(\mathrm{d}x, \mathrm{d}s), \end{aligned} \tag{2.4}$$

with $\|\Delta \widetilde{A}^T_t\| < c_A$ and $\|\Delta(A^T - \widetilde{A}^T)_t\| \geqslant c_A$, while the following assumptions are true:

(A1) The d-dimensional, bounded function γ has bounded variation on finite intervals, σ denotes a matrix-valued (symmetric), bounded and continuous function and W is a d-dimensional Wiener process independent of the jump part of A^T.

(A2) Let N^{A^T} be a Poisson random measure with intensity measure $\nu_{\cdot/2T} \otimes \lambda$, where

$$\nu_t(\mathrm{d}x) = \frac{c(t)}{\|x\|^{d+\alpha(t)}}\,\mathrm{d}x, \quad t \geqslant 0,$$

while $\alpha \in C^2(\mathbb{R}_+)$ and $c \in C(\mathbb{R}_+)$ are 1-periodic functions. For unique $a \in (0,1)$ and $b \in [0,1]$ it holds $\alpha_* := \alpha(a) < \alpha(t) < \alpha(b) =: \alpha^*$ for all $t \in [0,1] \setminus \{a,b\}$ and $0 < \alpha_* < \alpha^* < 2$ (Figure 1.8).

Remark 2.36. The additive process A^T given above admits the generating triplet $\left(\int_0^t \gamma(s)\,\mathrm{d}s, \int_0^t \sigma^2(s)\,\mathrm{d}s, \int_0^t \nu_{s/2T}(\cdot)\,\mathrm{d}s\right)$ (cf. Example 2.7), where ν_t denotes a Lévy measure of an $\alpha(t)$-stable Lévy process. The periodicity of the stability index $\alpha(t)$ of ν_t is a central property and the dependence on $2T$ – the length of one period – is the link to the topic of stochastic resonance.

Definition 2.37. *The compensated random measure* $N^{A^T}(\mathrm{d}x, \mathrm{d}s) - \nu_{s/2T}(\mathrm{d}x)\mathrm{d}s$ *is abbreviated by* $\widetilde{N}^{A^T}(\mathrm{d}x, \mathrm{d}s)$ *and* $\widetilde{A}^T$ *is called small jump part.*

Lemma 2.10 and condition (A1) imply that A^T is a semimartingale.

2.4.2 Existence, uniqueness and strong Markov property of solutions

In general existence and uniqueness results of stochastic differential equations demand global Lipschitz continuity.

Theorem 2.38. *Let Y be a d-dimensional semimartingale and $x_0 \in \mathbb{R}^d$. Let f be a matrix-valued function on $\mathbb{R}^d \times \mathbb{R}_+$ with Lipschitz continuous components. The stochastic differential equation of the form*

$$X_t = x_0 + \int_0^t f(X_{s-}, s)\,\mathrm{d}Y_s,$$

has a unique solution X which is a semimartingale.

Proof. The theorem immediately follows from Theorem 7 in Chapter 5 of [39]. □

Because the gradients ∇U and ∇V are not globally Lipschitz, a little detour is necessary.

Proposition 2.39. *The stochastic differential equation* (1.3)*, namely*

$$X_t^\varepsilon(x) = x - \int_0^t \nabla U\left(X_s^\varepsilon, \frac{s}{2T}\right)\mathrm{d}s + \varepsilon \int_0^t g\left(X_{s-}^\varepsilon\right)\mathrm{d}L_s, \quad t \geqslant 0,$$

has a unique solution which is a semimartingale.

Proof. The lack of a global Lipschitz constant of ∇U just guarantees the existence of a unique solution up to a certain explosion time and it remains to check the explosion time is almost surely infinite. The value of the parameter ε is irrevelant for this proof. Set $\varepsilon = 1$ and write

X instead of X^ε. Split the Lévy process L into a compound Poisson part P with jumps in $B_1^c(0)$ and $K = L - P$ with jump norm smaller than 1 (cf. Theorem 2.16). The sequence $(\tau_i)_{i\in\mathbb{N}}$ contains the jump times of P and $W_i = \Delta P_{\tau_i}$ the jumps. For the proof, we introduce a jump diffusion with bounded jumps.

Step 1: Let S be a stopping time and define $K_t^S := K_{S+t} - K_S$ which is a Lévy process with $K^S \stackrel{\mathrm{d}}{=} K$ (Theorem 32 in Chapter 1 of [39]). The jump diffusion with bounded jumps starting at S in a $\mathcal{F}_S$-measurable and finite state X_S satisfies the stochastic differential equation

$$X_{S,t}^K(X_S) = X_S - \int_0^t \nabla U\left(X_{S,r}^K, \frac{S+r}{2T}\right)\mathrm{d}r + \int_0^t g\left(X_{S,r-}^K\right)\mathrm{d}K_r^S, \quad t \geqslant 0.$$

The existence of a unique solution $(X_{S,t}^K(X_S))_{t\geqslant 0}$ and the strong Markov property of K^S (Lemma 2.12) imply that the solution X of (1.3) can be put together as

$$X_t(x) = \begin{cases} X_{0,t}^K(x), & t \in (0,\tau_1), \\ X_{0,\tau_1-}^K(x) + g\left(X_{0,\tau_1-}^K(x)\right)W_1, & t = \tau_1, \\ X_{\tau_1,t-\tau_1}^K(X_{\tau_1}(x)), & t \in (\tau_1,\tau_2), \\ X_{\tau_1,(\tau_2-\tau_1)-}^K(X_{\tau_1}(x)) + g\left(X_{\tau_1,(\tau_2-\tau_1)-}^K(X_{\tau_1}(x))\right)W_2, & t = \tau_2, \\ \dots . \end{cases}$$

Thus it remains to verify the existence and uniqueness of $(X_{S,t}^K(X_S))_{t\geqslant 0}$.

Step 2: For any $n \in \mathbb{N}$ define the continuous and bounded function $h_n : \mathbb{R}^d \to \mathbb{R}^d$ by

$$h_n(x) = \begin{cases} x, & \|x\| \leqslant n, \\ n\,\frac{x}{\|x\|}, & \text{otherwise}, \end{cases} \tag{2.5}$$

and consider the modified equation

$$X_{S,t}^{K,n}(X_S) = X_S - \int_0^t \nabla U\left(h_n\left(X_{S,r}^{K,n}\right), \frac{S+r}{2T}\right)\mathrm{d}r + \int_0^t g\left(X_{S,r-}^{K,n}\right)\mathrm{d}K_r^S, \quad t \geqslant 0$$

with Lipschitz continuous drift $\nabla U(h_n(\cdot), \frac{S+\cdot}{2T})$. For every $n \in \mathbb{N}$ a unique solution $(X_{S,t}^{K,n}(X_S))_{t\geqslant 0}$ exists (Theorem 2.38). Define the stopping times

$$\zeta_n := \inf\left\{r \geqslant 0 : \left\|X_{S,r}^{K,n}(X_S)\right\| \geqslant n - |||g|||\right\},$$

where $|||g|||$ denotes a matrix norm, which is an upper bound for the jump norm of the jump process $X_S^{K,n}$ due to $\|\Delta K_t^S\| \leqslant 1$. Then $X_{S,t}^{K,n}(X_S)$ and $X_{S,t}^{K,n+1}(X_S)$ coincide for $t < \zeta_n$. Obviously $(X_{S,t}^K(X_S))_{t\in[0,\lim_{n\to\infty}\zeta_n)}$ with $X_{S,t}^K(X_S) = X_{S,t}^{K,n}(X_S)$ for $t \in [0,\zeta_n)$ solves the original equation discussed in step 1.

Step 3: It remains to verify that the limit of ζ_n is almost surely infinite. Assume the contrary, thus there is a $z > 0$ with $\mathbb{P}(\lim_{n\to\infty}\zeta_n \leqslant z) > 0$. From the initial assumptions (U3) and (U4) we can derive the uniform growth of $U(x,t)$ as $\|x\|$ tends to infinity. Precisely we get

$$U(x,t) \geqslant \frac{c_U^*}{c_1}\|x\|^{1+c_U^*} - 1,$$

for $x \in \mathbb{R}^d \backslash O_{R^*}^U$ and $t \geqslant 0$ if $c_1 > 0$ denotes the uniform upper bound of $\|\nabla f(x,t)\|$ with $f(x,t) = \log(1+U(x,t))$. Due to $O_{R_1}^U \subseteq O_{R_2}^U$, $R_1 \leqslant R_2$ and $\bigcup_{R\geqslant 0} O_R^U = \mathbb{R}^d$ there is a $N_0 \in \mathbb{N}$ such that $O_{R^*}^U$ is contained in the ball $B_{N_0}(0)$. For all natural numbers $n \geqslant N_0$, $t \geqslant 0$, and $x \in B_n^c(0)$ this yields

$$\log(1+U(x,t)) \geqslant \log\left(\frac{c_U^*}{c_1}\, n^{1+c_U^*}\right) =: u_n.$$

Hence the estimate $f(X_{S,\zeta_n}^{K,n}(X_S), \frac{S+\zeta_n}{2T})) \geqslant u_{n-\lfloor |||g||| \rfloor} =: \bar{u}_n$ holds where $\lfloor \cdot \rfloor$ is the biggest previous integer. Applying Itô's formula (Theorem 2.30) to f, using the notation $U_t(x,t) = \frac{\partial}{\partial t} U(x,t)$, $K^S = (K^{S,1}, \dots, K^{S,d})$ and $f_{x_i x_j} = \frac{\partial^2}{\partial x_i \partial x_j} f$ yields

$$\begin{aligned}
&f\left(X_{S,t}^{K,n}, \frac{S+t}{2T}\right) \\
&= f\left(X_S, \frac{S}{2T}\right) + \frac{1}{2T}\int_0^t \frac{U_t\big(X_{S,r}^{K,n}, \frac{S+r}{2T}\big)}{1+U\big(X_{S,r}^{K,n}, \frac{S+r}{2T}\big)}\,\mathrm{d}r \\
&\quad - \int_0^t \frac{\|\nabla U\big(X_{S,r}^{K,n}, \frac{S+r}{2T}\big)\|^2}{1+U\big(X_{S,r}^{K,n}, \frac{S+r}{2T}\big)}\,\mathrm{d}r + \int_0^t \frac{\nabla U\big(X_{S,r-}^{K,n}, \frac{S+r}{2T}\big) g\big(X_{S,r-}^{K,n}\big)}{1+U\big(X_{S,r-}^{K,n}, \frac{S+r}{2T}\big)}\,\mathrm{d}K_r^S \\
&\quad + \frac{1}{2}\sum_{i,j,k,l=1}^{d} \int_0^t f_{x_i x_j}\left(X_{S,r}^{K,n}, \frac{S+r}{2T}\right) g_{ik}(X_{S,r}^{K,n})\, g_{jl}(X_{S,r}^{K,n})\,\mathrm{d}[K^{S,k}, K^{S,l}]_r^c \\
&\quad + \sum_{r\leqslant t}\left[f\left(X_{S,r}^{K,n}, \frac{S+r}{2T}\right) - f\left(X_{S,r-}^{K,n}, \frac{S+r}{2T}\right) - \nabla f\left(X_{S,r-}^{K,n}, \frac{S+r}{2T}\right)\Delta X_{S,r}^{K,n}\right]
\end{aligned}$$

for all $t \leqslant \zeta_n$ because then the drift term of the equation associated with $X_{S,t}^{K,n}$ equals to ∇U. Abbreviate all the summands in the formula for $f\left(X_{S,t}^{K,n}, \frac{S+t}{2T}\right)$ by $I_1(t), \dots, I_6(t)$ and realize that $I_3(t) = -\int_0^t \frac{\|\nabla U\|^2}{1+U}\,\mathrm{d}r$ is not positive. Now combine these assertions to estimate

$$\begin{aligned}
0 &< \mathbb{P}\left(\lim_{k\to\infty} \zeta_k \leqslant z\right) \\
&\leqslant \mathbb{P}\left(\zeta_n \leqslant z, f\left(X_{S,\zeta_n}^{K,n}(X_S), \frac{S+\zeta_n}{2T}\right) \geqslant \bar{u}_n\right) \\
&\leqslant \mathbb{P}\left(\zeta_n \leqslant z, \sum_{i=1,2,4,5,6} I_i(\zeta_n) \geqslant \bar{u}_n\right) \\
&\leqslant \mathbb{P}\left(\sup_{t\leqslant z} \sum_{i=1,2,4,5,6} I_i(t) \geqslant \bar{u}_n\right)
\end{aligned}$$

for all $n \geqslant \lfloor |||g||| \rfloor$. Thus it suffices to verify that

$$\sum_{i=1, i\neq 3}^{6} \mathbb{P}\left(\sup_{t\leqslant z} |I_i(t)| \geqslant \frac{\bar{u}_n}{5}\right)$$

is a null sequence as n tends to infinity. The finiteness of the random variable X_S immediately yields that the probability involving $I_1(t)$ tends to zero and from assumption (U4) we can derive

$$\mathbb{P}\left(\sup_{t\leqslant z}|I_2(t)| \geqslant \frac{\bar{u}_n}{5}\right) \leqslant \mathbb{P}\left(\frac{c_2\, z}{2T} \geqslant \bar{u}_n\right) = 0,$$

if $|U_t(x,t)| < c_2$ and n is large enough. For a short and clear treatment of the probabilities involving $I_5(t)$ and $I_4(t)$ omit the arguments of all functions. The Kunita-Watanabe inequality (Theorem 25 in Chapter 2 of [39]) induces

$$\begin{aligned}
&\frac{1}{2}\Big|\sum_{i,j,k,l=1}^{d}\int_0^t f_{x_ix_j}\, g_{ik}\, g_{jl}\,\mathrm{d}[K^{S,k},K^{S,l}]^c_r\Big| \\
&\leqslant \frac{1}{2}\sum_{i,j,k,l=1}^{d}\left(\int_0^t f^2_{x_ix_j}\, g^2_{ik}\, g^2_{jl}\,\mathrm{d}[K^{S,k}]^c_r\right)^{1/2}\left(\int_0^t f^2_{x_ix_j}\, g^2_{ik}\, g^2_{jl}\,\mathrm{d}[K^{S,j}]^c_r\right)^{1/2}.
\end{aligned}$$

Assumptions (U4) and (N2) imply the boundedness of the integrands. Since the continuous part of the Lévy process L with triplet (a,Σ,ν) is Gaussian we have $[K^{S,k}]^c_t \overset{\mathrm{d}}{=} [K^k]^c_t = \Sigma^2_{kk}t$ and thus we obtain the inequality $\sup_{t\leqslant z}|I_5(t)| \leqslant c_3 z$ for some $c_3 > 0$ but $\mathbb{P}(c_3 z \geqslant \bar{u}_n)$ vanishes for large n.

Let $U_{x_i}(x,t)$ be the partial derivative of U with respect to x_i and $M^S_r := K^S_r - r\mathbb{E}K_1$ is a martingale (Theorems 32 and 41 in Chapter 1 of [39]). The summand $I_4(t)$ equals to

$$\begin{aligned}
I_4(t) &= \sum_{j=1}^{d}\int_0^t \sum_{i=1}^{d}\frac{U_{x_i}\, g_{ij}}{1+U}\,\mathrm{d}K^{S,j}_r \\
&= \sum_{j=1}^{d}\int_0^t \sum_{i=1}^{d}\frac{U_{x_i}\, g_{ij}}{1+U}\,\mathrm{d}M^{S,j}_r + \int_0^t \sum_{i,j=1}^{d}\frac{U_{x_i}\, g_{ij}\,\mathbb{E}K^j_1}{1+U}\,\mathrm{d}r.
\end{aligned}$$

From the boundedness of g_{ij} and the first derivatives of $\log(1+U(x,t))$ given through assumption (N2) respectively (U4) and the finiteness of all moments of K^S because $\|\Delta K^S\| \leqslant 1$ (Theorem 34 in Chapter 1 of [39]) we derive

$$\mathbb{P}\left(\sup_{t\leqslant z}\left|\int_0^t \sum_{i,j=1}^{d}\frac{U_{x_i}\big(X^{K,n}_{S,r-},\frac{S+r}{2T}\big)g_{ij}\big(X^{K,n}_{S,r-}\big)\mathbb{E}K^j_1}{1+U\big(X^{K,n}_{S,r-},\frac{S+r}{2T}\big)}\,\mathrm{d}r\right| \geqslant \frac{\bar{u}_n}{10}\right) \leqslant \mathbb{P}\left(z\, c_4 \geqslant \bar{u}_n\right) = 0,$$

for some $c_4 > 0$ and all sufficiently large n. For brevity use the abbreviation $\nabla U g_{\cdot j} = \sum_{i=1}^d U_{x_i} g_{ij}$ to define the local martingale

$$Y^S_t = \sum_{j=1}^{d}\int_0^t \frac{\nabla U\, g_{\cdot j}}{1+U}\,\mathrm{d}M^{S,j}.$$

It is a square integrable martingale if and only if $\mathbb{E}[Y^S]_t < \infty$ for all $t \geqslant 0$ and then the equality $\mathbb{E}(Y^S_t)^2 = \mathbb{E}[Y^S]_t$ is true (Corollary 3 in Chapter 2 of [39]). The orthogonality of the continuous and the purely discontinuous part of the quadratic variation and the linearity in t of the continuous part cause the following

$$\left[Y^S\right]_t \leqslant c_5 t + \left[\sum_{j=1}^{d}\int_0^{\cdot}\frac{\nabla U\, g_{\cdot j}}{1+U}\,\mathrm{d}M^{S,j}_r\right]^d_t \leqslant c_6\left(t + \sum_{s\leqslant t}\|\Delta K^S_s\|^2\right),$$

for $c_5, c_6 > 0$, where the last inequality is a consequence of the boundedness of the integrand and $[M^S]^d = [K^S]^d$. The boundedness of $\mathbb{E}[Y^S]_t$ now harks back on an analysis of $\mathbb{E}\sum_{s\leqslant t}\|\Delta K_s^S\|^2$. The estimation

$$\mathbb{E}[Y^S]_t \leqslant tc_6\left(1+\int_{\|x\|\in(0,1)}\|x\|^2\,\mathrm{d}\nu(x)\right) =: tc_7,$$

Doob's martingale inequality (Theorem 3.8 in Chapter 1 of [34]), and the square integrability of Y^S now justify

$$\mathbb{P}\left(\sup_{t\leqslant z}|Y_t^S| \geqslant \frac{\bar{u}_n}{10}\right) \leqslant \frac{100}{\bar{u}_n^2}\,\mathbb{E}\left(Y_z^S\right)^2 = \frac{100}{\bar{u}_n^2}\,\mathbb{E}[Y^S]_z \leqslant \frac{100zc_7}{\bar{u}_n^2},$$

which tends to zero as $n\to\infty$. The term $I_6(t)$ can be treated with the Markov inequality

$$\mathbb{P}\left(\sup_{t\leqslant z}|I_6(t)| \geqslant \frac{\bar{u}_n}{5}\right) \leqslant \mathbb{P}\Big(c_8\sum_{t\leqslant z}\|\Delta K_t^S\|^2 \geqslant \bar{u}_n\Big) \leqslant \frac{c_8}{\bar{u}_n}\,\mathbb{E}\sum_{t\leqslant z}\|\Delta K_t^S\|^2,$$

which converges to zero as n diverges, for some $c_8 > 0$. All these estimates prove $\mathbb{P}\left(\lim_{k\to\infty}\zeta_k \leqslant z\right)$ is bounded from above by a null sequence and this yields the contradiction. □

Also allowing Marcus-type perturbations requires another existence and uniqueness result. In [32] a useful theorem is stated which again demands certain global Lipschitz conditions.

Theorem 2.40. ([32] Theorem 3.2.) *Assume the function f on $\mathbb{R}^d$ is matrix-valued, continuously differentiable, globally Lipschitz and $f'f$ is globally Lipschitz, too. Let X_0 be a $\mathcal{F}_0$-measurable random vector and Z denotes a semimartingale. Then the equation*

$$X_t = X_0 + \int_0^t f(X_s)\diamond \mathrm{d}Z_s,$$

has a unique solution which is a semimartingale.

Proposition 2.41. *The stochastic differential equation* (1.4)*, namely*

$$Z_t^\varepsilon(x) = x - \int_0^t \nabla U\left(Z_s^\varepsilon, \frac{s}{2T}\right)\mathrm{d}s + \varepsilon\int_0^t g(Z_s^\varepsilon)\diamond \mathrm{d}L_s, \quad t\geqslant 0,$$

has a unique solution that is a semimartingale.

The proof works analogous to the Itô case, but the notation is much more cumbersome.

Proof. **Step 1:** Again we deal with the truncated jump diffusion with bounded jumps $(Z_{S,t}^{K,n}(Z_S))_{t\geqslant 0}$ satisfying

$$Z_{S,t}^{K,n}(Z_S) = Z_S - \int_0^t \nabla U\left(h_n\left(Z_{S,r}^{K,n}(Z_S)\right), \frac{S+r}{2T}\right)\mathrm{d}r + \int_0^t g\left(Z_{S,r}^{K,n}(Z_S)\right)\diamond \mathrm{d}K_r^S,$$

where $h_n(x,t)$ is given in formula (2.5) in the proof of Proposition 2.39, K^S is defined at the beginning of the mentioned proof, $\varepsilon = 1$, S is a stopping time, and Z_S is $\mathcal{F}_S$-measurable and finite. On account of (N2) and

$$\varphi(t;x,y) = x + \int_0^t g(\varphi(r;x,y))y\,\mathrm{d}r$$

we obtain for some $C > 0$

$$\|\Delta Z_{S,t}^{K,n}\| \leqslant \left\|\varphi(1; Z_{S,t-}^{K,n}, \Delta K_t^S) - Z_{S,t-}^{K,n}\right\| \leqslant \left\| \int_0^1 g(\varphi(r; Z_{S,t-}^{K,n}, \Delta K_t^S))\Delta K_t^S \mathrm{d}r\right\| \leqslant C\|\Delta K_t^S\|$$

and that is why the jumps of $\Delta Z_{S,t}^{K,n}$ are bounded by some $0 < c_g < \infty$. The not truncated process $(Z_{S,t}^K(Z_S))_{t\geqslant 0}$ solves a similar equation with $\nabla U(h_n(\cdot), \frac{S+\cdot}{2T})$ replaced by ∇U. Define the stopping times

$$\zeta_n = \inf\left\{t \geqslant 0 : \left\|Z_{S,t}^{K,n}(Z_S)\right\| \geqslant n - c_g\right\}, \quad n \in \mathbb{N}.$$

This confirms that $Z_{S,t}^K(Z_S) = Z_{S,t}^{K,n}(Z_S)$ for $t \in [0,\zeta_n)$ and allows the splitting

$$Z_t(x) = \begin{cases} Z_{0,t}^K(x), & t \in (0,\tau_1), \\ \varphi(1; Z_{0,\tau_1-}^K(x), W_1), & t = \tau_1, \\ Z_{\tau_1,t-\tau_1}^K(Z_{\tau_1}(x)), & t \in (\tau_1,\tau_2), \\ \varphi(1; Z_{\tau_1,(\tau_2-\tau_1)-}^K(Z_{\tau_1}(x)), W_2), & t = \tau_2, \\ \dots, \end{cases}$$

where again $(\tau_n)_{n\in\mathbb{N}}$ are the jump times of $L - K$ and $(W_n)_{n\in\mathbb{N}}$ are the jumps. It suffices to prove the existence and uniqueness of $(Z_{S,t}^{K,n}(Z_S))_{t\geqslant 0}$. To embed the whole problem in the notation of Theorem 2.40 an enlargement of the state space from $\mathbb{R}^d$ to $\mathbb{R}^{d+1}$ is necessary. Define the column vectors $\bar{Z}_S = (Z_S, 0)^T$ and $\bar{K}_t^S = (K_t^S, t)^T$, the function $\bar{g}(\bar{x}) := g(\bar{x}_1, \dots, \bar{x}_d)$ for $\bar{x} = (\bar{x}_1, \dots, \bar{x}_{d+1}) \in \mathbb{R}^{d+1}$ and the $(d+1)\times(d+1)$-dimensional matrix

$$f(\bar{x}) = \begin{pmatrix} \bar{g}(\bar{x}) & -\nabla U(h_n(\bar{x}_1,\dots,\bar{x}_d), \bar{x}_{d+1}) \\ 0 & \frac{1}{2T} \end{pmatrix}, \quad \bar{x} \in \mathbb{R}^{d+1}.$$

The truncation of the space argument of ∇U and assumption (N2) guarantee the Lipschitz conditions of Theorem 2.40. According to that the stochastic differential equation

$$\begin{aligned} \bar{Z}_{S,t}^{K,n}(\bar{Z}_S) &= \bar{Z}_S + \int_0^t f\left(\bar{Z}_{S,r}^{K,n}(\bar{Z}_S)\right) \diamond \mathrm{d}\bar{K}_r^S \\ &= \bar{Z}_S + \int_0^t f\left(\bar{Z}_{S,r-}^{K,n}(\bar{Z}_S)\right) \mathrm{d}\bar{K}_r^S + \int_0^t f'\left(\bar{Z}_{S,r}^{K,n}(\bar{Z}_S)\right) f\left(\bar{Z}_{S,r}^{K,n}(\bar{Z}_S)\right) \mathrm{d}[\bar{K}^S, \bar{K}^S]_r^c \\ &\quad + \sum_{r\leqslant t} \left[\bar{\varphi}(1, \bar{Z}_{S,r-}^{K,n}(\bar{Z}_S), \Delta\bar{K}_r^S) - \bar{Z}_{S,r-}^{K,n}(\bar{Z}_S) - f(\bar{Z}_{S,r-}^{K,n}(\bar{Z}_S))\Delta\bar{K}_r^S\right] \end{aligned}$$

with

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\varphi}(t;\bar{x},\bar{y}) = f(\bar{\varphi}(t;\bar{x},\bar{y}))\bar{y}, \quad \bar{\varphi}(0;\bar{x},\bar{y}) = \bar{x}, \quad \bar{x},\bar{y} \in \mathbb{R}^{d+1},$$

admits a unique solution which is a semimartingale. Obviously $(\bar{Z}^{K,n}_{S,t}(\bar{Z}_S))_{t\geqslant 0}$ equals to $(Z^{K,n}_{S,t}(Z_S), \frac{t}{2T})^T_{t\geqslant 0}$, where $(Z^{K,n}_{S,t}(Z_S))_{t\geqslant 0}$ solves the stochastic differential equation stated in step 1. As in Proposition 2.39 the statement will follow from the almost sure limit of $\lim_{n\to\infty}\zeta_n = +\infty$.

Step 2: We repeat the arguments from the proof of Proposition 2.39 and apply the chain rule (Proposition 2.35). Define $\psi(\bar{x}) = \log(1+U(\bar{x}))$ for $\bar{x} \in \mathbb{R}^{d+1}$ and remember the definition of the constants u_n, $n \in \mathbb{N}$ in the last proof. We have $\psi(\bar{Z}^{K,n}_{S,\zeta_n}) \geqslant u_{n-\lfloor c_g\rfloor} =: \tilde{u}_n$. Inserting $\bar{Z}^{K,n}_{S,t}$ in ψ yields

$$\psi(\bar{Z}^{K,n}_{S,t}) = \psi(\bar{Z}_S) + \int_0^t \psi'(\bar{Z}^{K,n}_{S,r}) \diamond \mathrm{d}\bar{Z}^{K,n}_{S,r} = \psi(\bar{Z}_S) + \int_0^t \psi'(\bar{Z}^{K,n}_{S,r}) f(\bar{Z}^{K,n}_{S,r}) \diamond \mathrm{d}\bar{K}^S_r$$

with $\psi' = (\psi_{\bar{x}_1},\ldots,\psi_{\bar{x}_{d+1}})$. The last summand was defined in Definition 2.34 and

$$\begin{aligned}
&\int_0^t \psi'(\bar{Z}^{K,n}_{S,r}) f(\bar{Z}^{K,n}_{S,r}) \diamond \mathrm{d}\bar{K}^S_r \\
&= \int_0^t \psi'(\bar{Z}^{K,n}_{S,r}) f(\bar{Z}^{K,n}_{S,r})\,\mathrm{d}\bar{K}^S_r + \frac{1}{2}\operatorname{Trace}\left(\int_0^t \left(\psi'(\bar{Z}^{K,n}_{S,r}) f(\bar{Z}^{K,n}_{S,r})\right)' \mathrm{d}[\bar{K}^S,\bar{K}^S]^c_r\, f(\bar{Z}^{K,n}_{S,r})^T\right) \\
&\quad + \sum_{r\leqslant t}\left(\int_0^1 \left[\psi'(\bar{\varphi}(u;\bar{Z}^{K,n}_{S,r-},\Delta\bar{K}^S_r)) f(\bar{\varphi}(u;\bar{Z}^{K,n}_{S,r-},\Delta\bar{K}^S_r)) - \psi'(\bar{Z}^{K,n}_{S,r-}) f(\bar{Z}^{K,n}_{S,r-})\right]\mathrm{d}u\right)\Delta\bar{K}^S_r.
\end{aligned}$$

Abbreviate the three summands by $I_1(t)$, $I_2(t)$ and $I_3(t)$. A jump of $\psi(\bar{Z}^{K,n}_S)$ arises from the multiplication of $\Delta\bar{K}^S$ with $\int_0^1 \psi'(\bar{\varphi}(u))f(\bar{\varphi}(u))\,\mathrm{d}u$ which can be seen as averaging of $\psi' f$ over the integral curve connecting $\bar{Z}^K_{s,r-}$ with $\bar{Z}^K_{s,r}$. Remebering the definition f and $\bar{K}^S$ leads to the next equation for the term $I_1(t)$

$$\begin{aligned}
&\int_0^t \psi'(\bar{Z}^{K,n}_{S,r}) f(\bar{Z}^{K,n}_{S,r})\,\mathrm{d}\bar{K}^S_r \\
&= \int_0^t \nabla\psi(\bar{Z}^{K,n}_{S,r-})\bar{g}(\bar{Z}^{K,n}_{S,r-})\,\mathrm{d}K^S_r - \int_0^t \frac{\|\nabla U(\bar{Z}^{K,n}_{S,r})\|^2}{1+U(\bar{Z}^{K,n}_{S,r})}\,\mathrm{d}r + \int_0^t \frac{1}{2T}\,\psi_{\bar{x}_{d+1}}(\bar{Z}^{K,n}_{S,r})\,\mathrm{d}r \\
&=: I_{1,1}(t) + I_{1,2}(t) + I_{1,3}(t),
\end{aligned}$$

with $\nabla\psi = (\psi_{\bar{x}_1},\ldots,\psi_{\bar{x}_d})$. Because $I_{1,2}(t) \leqslant 0$, it suffices to verify that

$$\mathbb{P}\left(\psi(\bar{Z}_S) \geqslant \frac{\tilde{u}_n}{5}\right) + \sum_{i=1,3}\mathbb{P}\left(\sup_{t\leqslant z}|I_{1,i}(t)| \geqslant \frac{\tilde{u}_n}{5}\right) + \sum_{i=2,3}\mathbb{P}\left(\sup_{t\leqslant z}|I_i(t)| \geqslant \frac{\tilde{u}_n}{5}\right)$$

tends to zero as $n \to \infty$ to reach a contradiction. Obviously $\mathbb{P}\left(\psi(\bar{Z}_S) \geqslant \frac{\tilde{u}_n}{5}\right)$ is a null sequence. For the treatment of $I_{1,1}(t)$ we inspect the previous proof because the integrand is again bounded. Due to assumption (U4) the term $|I_{1,3}(t)|$ admits an upper bound linear in t. The summand involving the trace of a complicate matrix is very cumbersome but the last row and column of $[\bar{K}^S, \bar{K}^S]^c_t$ are filled with zeros which paves the way to use the arguments of the previous proof for this term. Note the equality of $\bar{\varphi}(u;\bar{x},\bar{y})$ and $(\varphi(u;x,y),x_{d+1})^T$ for $\bar{y}=(y,0)$ and $\bar{x}=(x,x_{d+1})$ with $x,y\in\mathbb{R}^d, x_{d+1}\in\mathbb{R}$. Then $I_3(t)$ equals to

$$I_3(t)=\sum_{r\leqslant t}\left(\int_0^1\left[\frac{\nabla U(\varphi(u),\frac{r}{2T})g(\varphi(u))}{1+U(\varphi(u),\frac{r}{2T})}-\frac{\nabla U(\varphi(0),\frac{r}{2T})g(\varphi(0))}{1+U(\varphi(0),\frac{r}{2T})}\right]\mathrm{d}u\right)\Delta K^S_r,$$

which can be bounded by $c_3\sum_{r\leqslant t}\|\Delta K^S_r\|^2$ for some $c_3>0$ because of assumptions (U4) and (N2). Attribute this proof to the previous one where we already dealed with $[K^S]^d$. □

Proposition 2.42. *The stochastic differential equation* (1.5)*, namely*

$$Y^\varepsilon_t(y)=y-\int_0^t\nabla V(Y^\varepsilon_s)\,\mathrm{d}s+\varepsilon A^T_t,$$

has a unique solution that is a semimartingale.

Proof. The proof is similar to the proof of Proposition 2.39. Because of the use of additive noise the jump part is much easier to handle. Apply Lemma 2.21 to treat the quadratic variation of the martingale part of the perturbation. □

The set of solutions of stochastic differential equations with respect to Lévy processes is strongly linked to Markov processes (Section 6.4.2. in [2], Section 6 of Chapter 5 in [39]). In our case a non-autonomous coefficient is involved and hence the dependence on the starting point and the starting time as well contradicts the strong Markov property of X^ε. Through an enlargement to (X^ε_t,t) we can immediately surmount this difficulty. The transformation of time inhomogeneous processes into homogeneous ones is frequently used. Among several one recent work applying such transformations is [8] in which Feller evolution systems are considered.

Proposition 2.43. *Assume X^ε solves the stochastic differential equation in* (1.3)*, then $(X^\varepsilon_t,t)_{t\geqslant 0}$ is a strong Markov process.*

Proof. Define the sequence $(\zeta_k)_{k\in\mathbb{N}_0}$ of increasing stopping times by $\zeta_0=0$ and

$$\zeta_k=\inf\{t\geqslant 0:\|L_t\|\geqslant k\}.$$

The equality $\lim_{k\to\infty}\zeta_k=\infty$ a.s. is valid. For all $n\in\mathbb{N}$ the Lévy process L admits the decomposition into independent Lévy processes $L^{<n}$ and $L^{\geqslant n}$, while $L^{\geqslant n}$ possesses the characteristic triplet $(0,0,\nu(\cdot\cap\{\|x\|\geqslant n\}))$ and $L^{<n}=L-L^{\geqslant n}$ is a process with bounded jumps. The equality $L_t=L^{<2k}_t$ holds for $t<\zeta_k$. Consider the stochastic differential equation

$$X^{\varepsilon,k}_t(x)=x-\int_0^t\nabla U\left(X^{\varepsilon,k}_s,\frac{s}{2T}\right)\mathrm{d}s+\varepsilon\int_0^t g\left(X^{\varepsilon,k}_{s-}\right)\mathrm{d}L^{<2k}_s,\ t\geqslant 0,$$

while the usual perturbation term L was replaced by $L^{<2k}$. Due to the boundedness of the jumps of $L^{<2k}$ a proof owing to the same arguments as at the beginning of this section verifies the existence and uniqueness of $(X_t^{\varepsilon,k})_{t\geqslant 0}$. The equations $X_t^{\varepsilon,k} = X_t^{\varepsilon,k+1}$ and $X_t^{\varepsilon} = X_t^{\varepsilon,k}$ are valid for $t < \zeta_k$. In the following passing the lines of Section 2 in [41] will finish the proof. Assume $(X_t^{\varepsilon,k}, t)_{t\geqslant 0}$ denotes a strong Markov process. Then (X_t^{ε}, t) belongs to the set of strong Markov processes if for all bounded stopping times τ, $A \in \mathcal{F}_\tau$, and nonnegative, bounded, and measurable functions f the succeeding equality holds

$$\int_A f(X_{\tau+t}^{\varepsilon}, \tau+t)\,\mathrm{d}\mathbb{P} = \int_A \mathbb{E}\left(f(X_{\tau+t}^{\varepsilon}, \tau+t)|(X_\tau^{\varepsilon}, \tau)\right)\mathrm{d}\mathbb{P}.$$

Applying the monotone convergence theorem twice and using $X_t^{\varepsilon} = X_t^{\varepsilon,k}$ for $t < \zeta_k$ and the strong Markov property of $(X_t^{\varepsilon,k}, t)_{t\geqslant 0}$ results in

$$\begin{aligned}
&\int_A f(X_{\tau+t}^{\varepsilon}, \tau+t)\,\mathrm{d}\mathbb{P} \\
&= \lim_{k\to\infty} \int_{A\cap\{\|L_r\|\leqslant k, r\in[0,\tau+t]\}} f(X_{\tau+t}^{\varepsilon,k}, \tau+t)\,\mathrm{d}\mathbb{P} \\
&= \lim_{k\to\infty} \int_{A\cap\{\|L_r\|\leqslant k, r\in[0,\tau]\}} \mathbb{E}\left(\mathbf{1}_{\{\|L_r\|\leqslant k, r\in[\tau,\tau+t]\}} f(X_{\tau+t}^{\varepsilon,k}, \tau+t)\Big|(X_\tau^{\varepsilon,k}, \tau)\right)\mathrm{d}\mathbb{P} \\
&= \int_A \mathbb{E}\left(f(X_{\tau+t}^{\varepsilon}, \tau+t)\big|(X_\tau^{\varepsilon}, \tau)\right)\mathrm{d}\mathbb{P}.
\end{aligned}$$

It remains to prove the strong Markov property of $\left(X_t^{\varepsilon,k}, t\right)_{t\geqslant 0}$. A result in [2] (Theorem 6.4.5) requires Lipschitz coefficients. Again truncate the space argument of ∇U to obtain the first condition. Define h_n as in formula (2.5) in the proof of Proposition 2.39. We consider

$$X_t^{\varepsilon,k,n}(x) = x - \int_0^t \nabla U\left(h_n\left(X_s^{\varepsilon,k,n}\right), \frac{s}{2T}\right)\mathrm{d}s + \varepsilon\int_0^t g\left(X_{s-}^{\varepsilon,k,n}\right)\mathrm{d}L_s^{<2k}, \quad t \geqslant 0.$$

Again the identity $X_t^{\varepsilon,k} = X_t^{\varepsilon,k,n} = X_t^{\varepsilon,k,n+1}$ is valid for $t < \sigma_n^k := \inf\left\{r \geqslant 0 : \|X_r^{\varepsilon,k,n}\| \geqslant n\right\}$. Thus at least $(X_t^{\varepsilon,k,n}, t)_{t\geqslant 0}$ is a strong Markov process. The inheritance to the process $\left(X_t^{\varepsilon,k}, t\right)_{t\geqslant 0}$ is again caused by the use of the method in [41] with $A \in \mathcal{F}_\tau$

$$\begin{aligned}
&\int_A f(X_{\tau+t}^{\varepsilon,k}, \tau+t)\,\mathrm{d}\mathbb{P} \\
&= \lim_{n\to\infty} \int_{A\cap\left\{\|X_r^{\varepsilon,k,n}\|\leqslant n, r\in[0,\tau]\right\}} \mathbb{E}\left(\mathbf{1}_{\left\{\|X_r^{\varepsilon,k,n}\|\leqslant n, r\in[\tau,\tau+t]\right\}} f(X_{\tau+t}^{\varepsilon,k,n}, \tau+t)\Big|(X_\tau^{\varepsilon,k,n}, \tau)\right)\mathrm{d}\mathbb{P} \\
&= \int_A \mathbb{E}\left(f(X_{\tau+t}^{\varepsilon,k}, \tau+t)\big|(X_\tau^{\varepsilon,k}, \tau)\right)\mathrm{d}\mathbb{P}
\end{aligned}$$

where τ is a bounded stopping time and f is nonnegative, bounded and measurable. □

Analogously the enlarged solution $(Z_t^{\varepsilon}, t)_{t\geqslant 0}$ can be treated.

Proposition 2.44. *Let Z^{ε} be the solution of* (1.4)*, then $(Z_t^{\varepsilon}, t)_{t\geqslant 0}$ is a strong Markov process.*

Proof. Kurtz, Pardoux and Protter provide us with an analogue of the just applied result of Protter for stochastic differential equations using Marcus integrals (Theorem 5.1 in [32]).

Their assertion requires an autonomous, Lipschitz coefficient f with $f'f$ Lipschitz and the perturbation process must be a Lévy process. As earlier we first cut off big jumps. Define the increasing stopping times $(\zeta_k)_{k\in\mathbb{N}_0}$ through $\zeta_0 = 0$ and

$$\zeta_k = \inf\{t \geqslant 0 : \|L_t\| \geqslant k\}$$

and remember the decomposition $L = L^{<2k} + L^{\geqslant 2k}$ to introduce $\bar{L}_t^{<2k} := (L_t^{<2k}, t)$. Define

$$f(\bar{x}_1, \ldots, \bar{x}_{d+1}) := \begin{pmatrix} \varepsilon g(\bar{x}_1, \ldots, \bar{x}_d) & -\nabla U\left(\bar{x}_1, \ldots, \bar{x}_d, \frac{\bar{x}_{d+1}}{2T}\right) \\ 0 & 1 \end{pmatrix}, \quad \bar{x} \in \mathbb{R}^{d+1}$$

and $\bar{x} = (x, 0)^T$. The stochastic differential equations

$$\bar{Z}_t^{\varepsilon,k}(\bar{x}) = \bar{x} + \int_0^t f(\bar{Z}_s^{\varepsilon,k}) \diamond \mathrm{d}\bar{L}_s^{<2k}, \quad k \in \mathbb{N},$$

have unique solutions that fulfill $\bar{Z}_t^{\varepsilon,k} = \bar{Z}_t^{\varepsilon,k+1}$ for $t < \zeta_k$. If $(\bar{Z}_t^{\varepsilon,k})_{t\geqslant 0}$ satisfies the strong Markov property, this again transfers to $(\bar{Z}_t^{\varepsilon})_{t\geqslant 0}$ by arguments of Samorodnitsky and Grigoriu ([41]). It remains to check that $(\bar{Z}_t^{\varepsilon,k})_{t\geqslant 0}$ really displays this property. Consider

$$\bar{Z}_t^{\varepsilon,k,n}(\bar{x}) = \bar{x} + \int_0^t f_n(\bar{Z}_s^{\varepsilon,k,n}) \diamond \mathrm{d}\bar{L}_s^{<2k},$$

where

$$f_n(\bar{x}_1, \ldots, \bar{x}_{d+1}) := \begin{pmatrix} \varepsilon g(\bar{x}_1, \ldots, \bar{x}_d) & -\nabla U\left(h_n\left(\bar{x}_1, \ldots, \bar{x}_d\right), \frac{\bar{x}_{d+1}}{2T}\right) \\ 0 & 1 \end{pmatrix}, \quad \bar{x} \in \mathbb{R}^{d+1}$$

with h_n defined in formula (2.5) in the proof of Proposition 2.39. This yields a unique solution $(\bar{Z}_t^{\varepsilon,k,n})_{t\geqslant 0} = (Z_t^{\varepsilon,k,n}, t)_{t\geqslant 0}$ that fulfills $\bar{Z}_t^{\varepsilon,k,n} = \bar{Z}_t^{\varepsilon,k,n+1}$ for t smaller than $\sigma_n^k := \inf\left\{t \geqslant 0 : \|Z_t^{\varepsilon,k,n}\| \geqslant n\right\}$. Theorem 5.1 in [32] implies the strong Markov property of $(\bar{Z}_t^{\varepsilon,k,n})_{t\geqslant 0}$. On account of [41] the process $(\bar{Z}_t^{\varepsilon,k})_{t\geqslant 0}$ shares this nice property with the twice truncated process. $\square$

Proposition 2.45. *Assume Y^ε solves the stochastic differential equation* (1.5), *then $(Y_t^\varepsilon, t)_{t\geqslant 0}$ is a strong Markov process.*

Proof. Similar to the proof of Theorem 6.4.5 in [2] we can verify that the solution $(Y_t^\varepsilon)_{t\geqslant 0}$ is Markov. Enlarging the state space by a time component guarantees that $(Y_t^\varepsilon, t)_{t\geqslant 0}$ is a strong Markov process. The same result can be also obtained by the arguments of Proposition 2.43. $\square$

2.5 Laplace's method

This nearly 200-year-old technique developed by Laplace serves for asymptotic evaluation of integrals depending on a parameter. Good references on this topic are [36] (Chapter 3, Section 7-9) and [46] (Chapter 2, Section 1).

Assume $\lambda > 0$ denotes a parameter, $f\colon \mathbb{R} \to \mathbb{R}_+$ and $g\colon \mathbb{R} \to \mathbb{R}_+$ are independent of λ. The approximation of

$$I(\lambda) := \int_a^b g(t)\,\mathrm{e}^{-\lambda f(t)}\,\mathrm{d}t$$

for large λ decisively depends on the position of the minimum of f in the possibly infinite interval (a, b) because the integrand shows a great peak near the minimum due to the exponential structure.

Assume g is continuous, f is twice differentiable and $f'(c) = 0$ for $c \in (a, b)$ and $f''(c) > 0$. Thus at c there is a minimum located and we assume this is the only one in (a, b). Replace $f(t)$ by its Taylor approximation

$$f(c) + \frac{1}{2} f''(c)(t - c)^2$$

and $g(x)$ by $g(c) \neq 0$, extend the integral limits to infinity and use the Poisson integral to estimate

$$\begin{aligned} I(\lambda) &\approx g(c)\,\mathrm{e}^{-\lambda f(c)} \int_a^b \mathrm{e}^{-\lambda f''(c)(t-c)^2/2}\,\mathrm{d}t \\ &\approx g(c)\,\mathrm{e}^{-\lambda f(c)} \int_{-\infty}^{\infty} \mathrm{e}^{-\lambda f''(c)(t-c)^2/2}\,\mathrm{d}t \\ &= g(c)\,\mathrm{e}^{-\lambda f(c)} \frac{\sqrt{2\pi}}{\sqrt{\lambda f''(c)}}. \end{aligned}$$

This gives a good approximation through very simple tools.

If f attains its smallest value at an end point of the interval (a, b) for example at a but $f'(a) > 0$ and $g(a) \neq 0$, then we use the first order approximation to obtain

$$\begin{aligned} I(\lambda) &\approx g(a)\,\mathrm{e}^{-\lambda f(a)} \int_a^b \mathrm{e}^{-\lambda f'(c)(t-a)}\,\mathrm{d}t \\ &\approx g(a)\,\mathrm{e}^{-\lambda f(a)} \int_a^{\infty} \mathrm{e}^{-\lambda f'(a)(t-a)}\,\mathrm{d}t \\ &= \frac{g(a)\,\mathrm{e}^{-\lambda f(a)}}{\lambda f'(a)}. \end{aligned}$$

In general it suffices to consider the case with $f(a) < f(t)$ for all $t \in (a, b]$ because otherwise intervals can be splitted such that this condition is fulfilled. The following theorem is taken from [36] (Theorem 7.1 in Chapter 3).

Proposition 2.46. *Let $f\colon \mathbb{R} \to \mathbb{R}_+$ be differentiable and $g\colon \mathbb{R} \to \mathbb{R}_+$. Assume $a < b \leqslant \infty$ with $f(a) < f(t)$ for all $t \in (a, b)$ while the minimum is attained only at a and $f'(t)$ and $g(t)$ are continuous in the neighbourhood of a except possibly at a. As t approaches a from the right we demand*

$$f(t) - f(a) = F(t-a)^x + \mathrm{O}((t-a)^{x+1}), \qquad g(t) = G(t-a)^{y-1} + \mathrm{O}((t-a)^y),$$

where the first relations are differentiable with $x, y > 0$ and $F, G \in \mathbb{C} \setminus \{0\}$. Then it holds

$$I(\lambda) = \frac{G}{x}\,\Gamma\left(\frac{y}{x}\right) \frac{\mathrm{e}^{-\lambda f(a)}}{(F\lambda)^{y/x}} + \mathrm{O}\left(\lambda^{-(y+1)/x}\right)$$

as λ tends to infinity.

In the present work the succeeding version of this method will be important.

Lemma 2.47. *Let $\alpha\colon [0,1] \to \mathbb{R}_+$ be two times differentiable with $\min_{t\in[0,1]} \alpha(t) = \alpha(a) =: \alpha_*$ for some unique $a \in (0,1)$ and $g\colon \mathbb{R} \to \mathbb{R}_+$ be continuous. Then it holds*

(i)
$$\lim_{\varepsilon\to 0} \sqrt{|\log\varepsilon|}\frac{\sqrt{\alpha''(a)}}{\sqrt{2\pi}\, g(a)} \int_0^1 \varepsilon^{\alpha(t)-\alpha_*} g(t)\,\mathrm{d}t = 1,$$

(ii)
$$\lim_{\varepsilon\to 0} \sqrt{|\log\varepsilon|}\frac{\sqrt{\alpha''(a)}}{\sqrt{2\pi}\, g(a)} \int_0^a \varepsilon^{\alpha(t)-\alpha_*} g(t)\,\mathrm{d}t = \frac{1}{2}.$$

Proof. At first we prove that 1 is a lower bound for the limit in (i). For all $\delta > 0$ there exists a $\xi > 0$ such that $|t-a| < \xi$ yields

$$\alpha''(t) \leqslant \alpha''\,(a) + \delta.$$

From the Taylor formula we can deduce for t within the ξ-neighbourhood of a

$$\alpha(t) - \alpha_* \leqslant \big(\alpha''\,(a) + \delta\big)\,(t-a)^2\,\frac{1}{2}.$$

Analogously there is a $\tilde{\xi} > 0$ that guarantees $g(t) \geqslant g\,(a) - \delta$ for $|t-a| < \tilde{\xi}$. Define $\bar{\xi} = \xi \wedge \tilde{\xi} \wedge a \wedge (1-a)$ and estimate

$$\begin{aligned}\int_0^1 \varepsilon^{\alpha(t)-\alpha_*} g(t)\,\mathrm{d}t &\geqslant \int_{a-\bar{\xi}}^{a+\bar{\xi}} g(t)\,\mathrm{e}^{-(\alpha(t)-\alpha_*)|\log\varepsilon|}\,\mathrm{d}t\\ &\geqslant (g\,(a)-\delta)\int_{a-\bar{\xi}}^{a+\bar{\xi}} \mathrm{e}^{-(\alpha''(a)+\delta)|\log\varepsilon|(t-a)^2/2}\,\mathrm{d}t.\end{aligned}$$

Substituting $u = \sqrt{(\alpha''(a)+\delta)|\log\varepsilon|}(t-a)$ yields the lower bound

$$\int_0^1 \varepsilon^{\alpha(t)-\alpha_*} g(t)\,\mathrm{d}t \geqslant \frac{(g\,(a)-\delta)\sqrt{2\pi}}{\sqrt{(\alpha''(a)+\delta)|\log\varepsilon|}} \int_{-\bar{\xi}\sqrt{(\alpha''(a)+\delta)|\log\varepsilon|}}^{\bar{\xi}\sqrt{(\alpha''(a)+\delta)|\log\varepsilon|}} \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-u^2/2}\,\mathrm{d}u.$$

In the limit as ε tends to zero the integral multiplied with an ε-dependent prefactor is bigger or equal to 1

$$\lim_{\varepsilon\to 0} \frac{\sqrt{\alpha''(a)|\log\varepsilon|}}{\sqrt{2\pi}g(a)} \int_0^1 \varepsilon^{\alpha(t)-\alpha_*} g(t)\mathrm{d}t \geqslant \frac{(g\,(a)-\delta)\sqrt{\alpha''(a)}}{g(a)\sqrt{\alpha''(a)+\delta}} \xrightarrow{\delta\to 0} 1.$$

Now concentrate on the upper bound and choose $\delta_0 > 0$ such that $-\alpha''(a) + \delta_0 < 0$. Then for all $\delta \in [0, \delta_0)$ there is a $\xi > 0$ such that

$$\alpha(t) - \alpha_* \geqslant \big(\alpha''\,(a) - \delta\big)\,(t-a)^2\,\frac{1}{2}$$

for all t with $|t-a|<\xi$. Again there is a $\tilde{\xi}>0$ with $g(t)\leqslant g\,(a)+\delta$ for $|t-a|<\tilde{\xi}$. Define $\bar{\xi}=\xi\wedge\tilde{\xi}\wedge a\wedge(1-a)$. Then $\beta>0$ exists such that $\alpha(t)$ is bigger than $\alpha_*+\beta$ for all t outside of the $\bar{\xi}$-neighbourhood of a. Splitting the interval $[0,1]$ into $[0,a-\bar{\xi}]$, $[a-\bar{\xi},a+\bar{\xi}]$ and $[a+\bar{\xi},1]$ suggests the upper bound

$$\begin{aligned}
&\int_0^1 \varepsilon^{\alpha(t)-\alpha_*} g(t)\,\mathrm{d}t\\
&\leqslant \int_0^{a-\bar{\xi}} g(t)\,\mathrm{e}^{-\beta|\log\varepsilon|}\,\mathrm{d}t+\int_{a-\bar{\xi}}^{a+\bar{\xi}}(g\,(a)+\delta)\,\mathrm{e}^{-(\alpha''(a)-\delta)|\log\varepsilon|(t-a)^2/2}\,\mathrm{d}t+\int_{a+\bar{\xi}}^1 g(t)\,\mathrm{e}^{-\beta|\log\varepsilon|}\,\mathrm{d}t\\
&\leqslant \varepsilon^{\beta}\max_{t\in[0,1]} g(t)+\frac{(g\,(a)+\delta)\sqrt{2\pi}}{\sqrt{(\alpha''(a)-\delta)|\log\varepsilon|}}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-u^2/2}\,\mathrm{d}u,
\end{aligned}$$

while a similar substitution $u=\sqrt{(\alpha''(a)-\delta)|\log\varepsilon|}(t-a)$ was used as before. This justifies the asymptotic upper bound

$$\lim_{\varepsilon\to0}\frac{\sqrt{\alpha''(a)|\log\varepsilon|}}{\sqrt{2\pi}g(a)}\int_0^1\varepsilon^{\alpha(t)-\alpha_*}g(t)\,\mathrm{d}t\leqslant\frac{(g\,(a)+\delta)\sqrt{\alpha''(a)}}{g(a)\sqrt{\alpha''(a)-\delta}}\xrightarrow{\delta\to0}1,$$

because $\varepsilon^{\beta}|\log\varepsilon|^{1/2}$ tends to zero. The proof of (ii) is analogous. The limit $\frac{1}{2}$ is obtained due to

$$\int_{a-\bar{\xi}}^{a}\mathrm{e}^{-(\alpha''(a)+\mathrm{O}(\delta))|\log\varepsilon|(t-a)^2/2}\,\mathrm{d}t=\int_{-\bar{\xi}\sqrt{(\alpha''(a)+\mathrm{O}(\delta))|\log\varepsilon|}}^{0}\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-u^2/2}\,\mathrm{d}u\xrightarrow{\varepsilon\to0}\frac{1}{2}.$$

□

Chapter 3

Stochastic resonance for stochastic differential equations driven by multiplicative Lévy noise with heavy tails

Assume the potential U satisfies the conditions (U1)-(U4) stated in Section 2.1 and assumptions (N1), (N2) and (T) presented in Subsection 2.4.1 hold true. For simplicity we assume that through the geometry of U and the definition of the Radon measure μ as the vague limit of the jump measure of L a transition from the well corresponding to $m_-(t)$ to the well containing $m_+(t)$ is most likely to occur at $(2k+1)T$, $k \in \mathbb{N}_0$ and a transition back is very probable at $2kT$, $k \in \mathbb{N}$. A rigorous definition will follow.

At first a multiplicative Lévy noise serves as a perturbation term. Namely, we shall consider $X^\varepsilon = (X_t^\varepsilon)_{t \geqslant 0}$ being the solution of the Itô stochastic differential equation driven by a regularly varying Lévy process L with index $-\alpha < 0$

$$X_t^\varepsilon(x) = x - \int_0^t \nabla U\left(X_s^\varepsilon, \frac{s}{2T}\right) \mathrm{d}s + \varepsilon \int_0^t g\left(X_{s-}^\varepsilon\right) \mathrm{d}L_s, \quad t \geqslant 0. \tag{3.1}$$

Also perturbations by a multiplicative noise in the Marcus sense are of interest, namely a process $Z^\varepsilon = (Z_t^\varepsilon)_{t \geqslant 0}$ being the solution of

$$Z_t^\varepsilon(x) = x - \int_0^t \nabla U\left(Z_s^\varepsilon, \frac{s}{2T}\right) \mathrm{d}s + \varepsilon \int_0^t g(Z_s^\varepsilon) \diamond \mathrm{d}L_s, \quad t \geqslant 0. \tag{3.2}$$

In the first section of this chapter we consider approximative two-state Markov chains living on the well minima with periodic transition rates of order ε^α to become familiar with the dynamics of the jump diffusions X^ε and Z^ε. We calculate the probability of a jump of the Markov chain within a small time interval surrounding the point where the transition is most likely. Afterwards the deviation of the small jump process of X^ε respectively Z^ε from the deterministic solution of $\frac{\mathrm{d}}{\mathrm{d}x}x(t) = -\nabla U\left(x(t), \frac{t}{2T}\right)$ is examined. The last section deals with the application of the same probabilistic quality measure of tuning of the Markov chains to the jump diffusions.

3.1 Approximation by two-state Markov chains

Before a detailed analysis of the solution $(X_t^\varepsilon)_{t\geqslant 0}$ of the stochastic differential equation (3.1) is possible, it seems reasonable to discuss the stochastic dynamics of reduced models. The process $(X_t^\varepsilon)_{t\geqslant 0}$ describes the perturbed movement of a particle within a periodically changing double-well potential. Since the intra-well fluctuations should be irrevelant for stochastic resonance, we now only consider time-continuous Markov chains that admit two values representing the two wells.

3.1.1 Piecewise constant infinitesimal generator

Define the time-inhomogeneous Markov chain $K = (K_t^\varepsilon)_{t\geqslant 0}$ on the state space $\{-1, 1\}$ with piecewise constant infinitesimal generator $Q^K = (Q_t^K)_{t\geqslant 0}$ given by

$$Q_t^K = \begin{cases} \begin{pmatrix} -\varphi_K & \varphi_K \\ \psi_K & -\psi_K \end{pmatrix} =: \mathcal{Q}_1, & t \in [2kT, (2k+1)T),\ k \in \mathbb{N}_0, \\ \begin{pmatrix} -\psi_K & \psi_K \\ \varphi_K & -\varphi_K \end{pmatrix} =: \mathcal{Q}_2, & t \in [(2k+1)T, (2k+2)T),\ k \in \mathbb{N}_0. \end{cases}$$

In this case the source of periodicity is the periodic switching between two asymmetric potential shapes U_1 and U_2, see in Figure 3.1. The left wells of U_1 and U_2 are identified with -1 and the right ones with 1. The roles of the transition rates φ_K and ψ_K switch after time T because an exit from the well belonging to $m_-(t)$ of the jump diffusion is favoured at time points $(2k+1)T$, $k \in \mathbb{N}_0$ and from the other one at $2kT$, $k \in \mathbb{N}$.
We consider the following appropriate regularly varying transition rates φ_K and ψ_K

$$\varphi_K = \varphi_K(\varepsilon) = \frac{\varepsilon^\alpha l(\varepsilon^{-1})}{a^\alpha}, \qquad \psi_K = \psi_K(\varepsilon) = \frac{\varepsilon^\alpha l(\varepsilon^{-1})}{b^\alpha},$$

where $\alpha > 0$ and $0 < b < a$ are fixed and l denotes a slowly varying function.
Besides [24], especially [37] suggests the computation of the invariant measure of the extended, time-scaled, and therefore homogeneous process

$$\widetilde{K}_t^\varepsilon := (K_{2Tt}^\varepsilon, t \bmod 2T)_{t\geqslant 0},$$

which then paves the way to compute the spectral power amplification, energy, relative entropy and other quality measures of tuning. Fortunately it is possible to adapt the procedure of Section 1 in Chapter 4 of [37] to obtain the invariant measure $\nu = (\nu^+, \nu^-)$ of $\widetilde{K}^\varepsilon$ on $[0, 1]$.

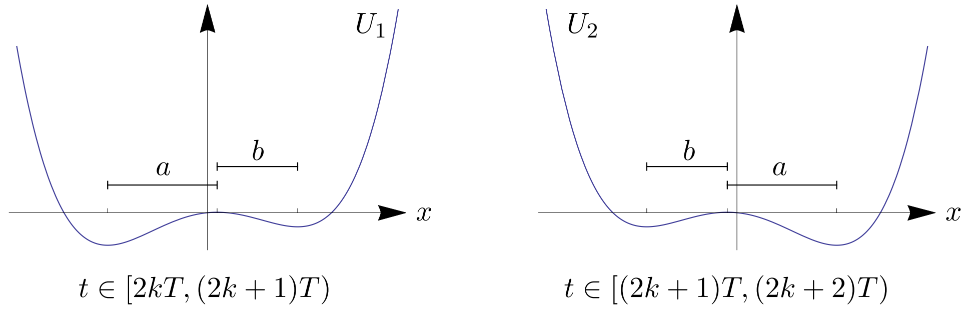


Figure 3.1: For $t \in [2kT, (2k+1)T)$ the transition rate from -1 to 1 is the inverse of the mean exit time of a jump diffusion of the left/bigger well of U_1 with barrier distance a. A transition from 1 to -1 corresponds to an exit of the right/smaller well with barrier distance b. If $t \in [(2k+1)T, (2k+2)T)$, the distances of the obstacle switch according to the picture of U_2 and the left well becomes the smaller one.

The value $\nu^{\pm}(t)$ gives the probability that K^{ε}_{2Tt} is ± 1 and it equals to

$$\nu^-(t) = \frac{b^{-\alpha}}{b^{-\alpha}+a^{-\alpha}} + \frac{(a^{-\alpha}-b^{-\alpha})\cdot \mathrm{e}^{-2T\,t\,\varepsilon^{\alpha} l(\varepsilon^{-1})(a^{-\alpha}+b^{-\alpha})}}{(a^{-\alpha}+b^{-\alpha})\cdot\big(1+\mathrm{e}^{-T\,\varepsilon^{\alpha} l(\varepsilon^{-1})(a^{-\alpha}+b^{-\alpha})}\big)}, \quad t \in \Big[0, \frac{1}{2}\Big),$$

$$\nu^+(t) = \frac{a^{-\alpha}}{b^{-\alpha}+a^{-\alpha}} - \frac{(a^{-\alpha}-b^{-\alpha})\cdot \mathrm{e}^{-2T\,t\,\varepsilon^{\alpha} l(\varepsilon^{-1})(a^{-\alpha}+b^{-\alpha})}}{(a^{-\alpha}+b^{-\alpha})\cdot\big(1+\mathrm{e}^{-T\,\varepsilon^{\alpha} l(\varepsilon^{-1})(a^{-\alpha}+b^{-\alpha})}\big)}, \quad t \in \Big[0, \frac{1}{2}\Big).$$

The equation $Q_t = \mathcal{Q}_2$ for $t \in \big[\frac{1}{2}, 1\big)$ leads to the symmetry conditions

$$\nu^-\left(t+\frac{1}{2}\right) = \nu^+(t), \quad t \in \Big[0, \frac{1}{2}\Big), \qquad \nu^+\left(t+\frac{1}{2}\right) = \nu^-(t), \quad t \in \Big[0, \frac{1}{2}\Big).$$

A frequently used tool to reveal underlying periodicities is the spectral power amplification η defined up to a constant pre-factor as

$$\eta(\varepsilon, T) = \left| \int_0^1 \mathbb{E}_{\nu} K^{\varepsilon}_{2Tt}\, \mathrm{e}^{2\pi \mathrm{i} t}\, \mathrm{d}t \right|^2$$

where $\mathbb{E}_{\nu}$ is the expectation with respect to ν. The amplification of a periodic perturbation usually is visualized through a well pronounced spectral peak at the corresponding frequency for the noisy output signal. For Gaussian diffusions this quantity emerges as suitable since a unique maximum exists ([37] Proposition 4.2.2). In what extent this applies for the Lévy-driven diffusion is analysed below. With the help of some explicit calculations, we get that the spectral power amplification equals to

$$\eta(\varepsilon, T) = \frac{T^2\,\varepsilon^{2\alpha} l(\varepsilon^{-1})^2 (a^{-\alpha}-b^{-\alpha})^2}{T^2\,\varepsilon^{2\alpha} l(\varepsilon^{-1})^2 (a^{-\alpha}+b^{-\alpha})^2 + \pi^2}.$$

Maximization in ε yields the optimal noise level for an amplification. Assume for simplicity that $l \equiv 1$ as in the α-stable case. Then it follows

$$\frac{\partial}{\partial \varepsilon}\eta(\varepsilon, T) = \frac{2\,\alpha\,(a^{-\alpha}-b^{-\alpha})^2\,\varepsilon^{2\alpha-1}\,\pi^2\,T^2}{(\pi^2 + (a^{-\alpha}+b^{-\alpha})^2\,\varepsilon^{2\alpha}\,T^2)^2} \;>\; 0,$$

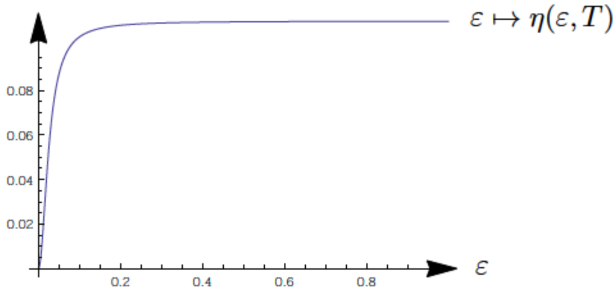


Figure 3.2: The function $\varepsilon \mapsto \eta(\varepsilon, T)$ is increasing if $l \equiv 1$, $a = 5$, $b = 1$, $\alpha = 1$ and $T = 100$.

and thus the function $\varepsilon \mapsto \eta(\varepsilon, T)$ is strictly increasing. The monotonicity makes the existence of a maximum impossible. Unfortunately the structure of the transition rates together with the plausible choice of the time scale

$$T(\varepsilon) = \varepsilon^{\alpha} l(\varepsilon^{-1})$$

immediately kill the dependence of the quality measure of tuning on the noise amplitude ε. Although it is touted as the universal quantity in many applications of stochastic resonance, it is not suitable here. The relevant time scales are of the same polynomial order and thus too similar.
Other tools which also use the invariant measure but do not originate from spectral theory, like the relative entropy ([37] Chapter 4, Section 7), also fail to detect the underlying periodicity. The relative entropy is a measure of disorder and chaos and is frequently applied in information theory. A minimum of the relative entropy reveals the frequency which causes the most deterministic behaviour. But here the entropy is monotonically decreasing in ε.

The lack of suitable quantities of tuning made a search for other periodicity discovering measures necessary. In the paper [19] Herrmann and Imkeller presented a totally different approach called "intensity of the first peak". The random output signal $K = (K_t^{\varepsilon})_{t \geqslant 0}$ looks quite periodic if K switches from ± 1 to ∓ 1 in a small time interval around $t = kT$. The concept of Herrmann and Imkeller rests on maximizing the distribution of jump times around multiples of T (especially T itself).

Definition 3.1. *Let $T_0 = 0$ and T_n, $n \in \mathbb{N}$, be the n-th jump time of K from -1 to 1 or vice versa. Then $\tau_n = \frac{T_n}{T}$, $n \in \mathbb{N}_0$, denotes the normalized jump time.*

The behaviour of τ_n only depends on the last normalized jump time τ_{n-1} and the value $K^{\varepsilon}_{T\tau_{n-1}}$ of the Markov chain. The value $K^{\varepsilon}_{T\tau_{n-1}}$ is uniquely determined by the value of $(-1)^{\lfloor \tau_{n-1} \rfloor + n}$. If $(-1)^{\lfloor \tau_{n-1} \rfloor + n}$ is equal to 1 either the n-th jump goes from 1 to -1 with the Q-matrix $\mathcal{Q}_1$ or the Markov chain K moves from -1 to 1 under the regime of $\mathcal{Q}_2$. All in all this means the recent transition rate is ψ_K which can be interpreted as the exit from the smaller potential

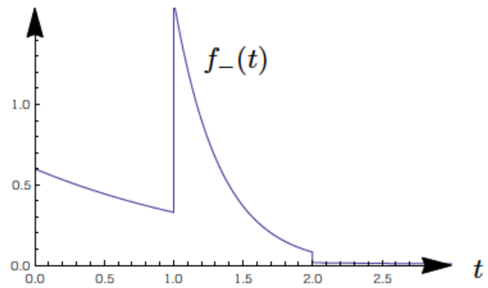


Figure 3.3: The density of τ_1 given $K_0 = -1$ with $l \equiv 1$, $a = 5$, $b = 1$, $\alpha = 1$ and $T = \frac{(a+b)/2}{\varepsilon}$.

well. Therefore $(-1)^{\lfloor \tau_{n-1} \rfloor + n} = -1$ stands for an upcoming exit from the bigger well with transition rate φ_K.

Lemma 3.2. *The law of the n-th normalized jump time τ_n, $n \in \mathbb{N}$, given $\tau_{n-1} = s$ and $(-1)^{\lfloor s \rfloor + n} = -1$ respectively 1 admits the density*

$$f_-(t) = r_0 T\, \mathrm{e}^{-r_0 T(t-s)}\, \mathbf{1}_{[s, \lfloor s \rfloor + 1)}(t) + \mathrm{e}^{-r_0 T(1+\lfloor s \rfloor - s)} \sum_{k=0}^{\infty} r_{k+1} T\, \mathbf{1}_{I_k}(t)\, \mathrm{e}^{-\sum_{i=1}^{k} r_i T - r_{k+1} T(t - \lfloor t \rfloor)},$$

for $t \geqslant s$, respectively

$$f_+(t) = r_1 T\, \mathrm{e}^{-r_1 T(t-s)}\, \mathbf{1}_{[s, \lfloor s \rfloor + 1)}(t) + \mathrm{e}^{-r_1 T(1+\lfloor s \rfloor - s)} \sum_{k=0}^{\infty} r_k T\, \mathbf{1}_{I_k}(t)\, \mathrm{e}^{-\sum_{i=1}^{k} r_{i-1} T - r_k T(t - \lfloor t \rfloor)},$$

for $t \geqslant s$ with

$$I_k = \{t \geqslant 0 : t \in [\lfloor s \rfloor + k + 1, \lfloor s \rfloor + k + 2)\},$$

$$r_k = r_k(\varepsilon) = \begin{cases} \varphi_K, & k \text{ even} \\ \psi_K, & k \text{ odd} \end{cases} = \begin{cases} \frac{\varepsilon^\alpha l(\varepsilon^{-1})}{a^\alpha}, & k \text{ even} \\ \frac{\varepsilon^\alpha l(\varepsilon^{-1})}{b^\alpha}, & k \text{ odd}, \end{cases}$$

for $k \in \mathbb{N}_0$.

Proof. Markov chains admit exponentially distributed inter-jump times and the infinitesimal generator of K periodically varies. This immediately gives the result (Lemma 1 in Chapter 1 of [19]). □

Due to the ε-dependence of r_k, $k \in \mathbb{N}_0$, the densities $f_\pm$ (example in Figure 3.3) and hence the jump behaviour of K change in the small noise limit. If $T = T(\varepsilon)$, the asymptotic law of jump times of K is determined by the limit of $r_k T$, $k = 0, 1$, as ε tends to zero. If this limit equals to zero, the weak limit of the law of the next jump time of K is the null measure and therefore reveals that time scales T of order $\varepsilon^{-\beta}$ with $\beta < \alpha$ are much too small to examine any jump. Contrary to that, time scales with $\beta > \alpha$ guarantee the observation of jumps, which in the mean occur at scale $\varepsilon^{-\alpha}$, but the distinction of single jumps is impossible in such a coarse time scale. All jumps are compressed to an instantaneous jump. For a rigorous treatment of the asymptotic behaviour of K we include the following lemma that applies the technique Herrmann and Imkeller used for the Gaussian case (Section 1.2 of [19]).

Lemma 3.3. *The conditional Laplace transform of τ_n given $\tau_{n-1} = s \geqslant 0$ and $(-1)^{\lfloor s \rfloor + n} = -1$ is*

$$\begin{aligned} L_-(x) = {} & \frac{r_0 T}{r_0 T + x} \left(\mathrm{e}^{-xs} - \mathrm{e}^{-x(\lfloor s \rfloor + 1) - r_0 T(\lfloor s \rfloor + 1 - s)} \right) \\ & + \mathrm{e}^{-r_0 T(\lfloor s \rfloor + 1 - s)} \sum_{k=0}^{\infty} \frac{r_{k+1} T}{r_{k+1} T + x} \, \mathrm{e}^{-\sum_{i=1}^{k} r_i T - x(\lfloor s \rfloor + k + 1)} \left(1 - \mathrm{e}^{-r_{k+1} T - x}\right) \end{aligned}$$

with $r_{2k} = \varphi_K$ and $r_{2k+1} = \psi_K$ for $k \in \mathbb{N}_0$ as in Lemma 3.2.

(i) If $\lim_{\varepsilon \to 0} r_k(\varepsilon) T(\varepsilon) = 0$ for $k = 0, 1$, the weak limit of the conditional law of τ_n is the null measure.

(ii) If $\lim_{\varepsilon \to 0} r_k(\varepsilon) T(\varepsilon) = \infty$ for $k = 0, 1$, the conditional law of τ_n weakly tends to the Dirac measure in the starting time s.

(iii) Assume

$$0 < \lim_{\varepsilon \to 0} r_0(\varepsilon) T(\varepsilon) =: c_0 < c_1 := \lim_{\varepsilon \to 0} r_1(\varepsilon) T(\varepsilon) < \infty.$$

If an exponentially distributed random variable $E_1 \sim \mathrm{Exp}(c_0)$ attains a value smaller than $\lfloor s \rfloor + 1 - s$ the conditional law of τ_n equals to an exponential distribution on the interval $[s, \lfloor s \rfloor + 1)$ represented by the random variable $E_2 \sim \mathrm{Exp}([s, \lfloor s \rfloor + 1), c_0)$ which is independent of E_1. For values bigger than $\lfloor s \rfloor + 1 - s$ the conditional law of τ_n equals to the law of a random variable Z. The following case differentiation holds:

$$Z = \begin{cases} X + \lfloor s \rfloor + 1, & G \text{ even}, \\ Y + \lfloor s \rfloor + 1, & G \text{ odd}, \end{cases}$$

while the random variables E_1, $X \sim \mathrm{Exp}([G, G+1), c_1)$, $Y \sim \mathrm{Exp}([G, G+1), c_0)$ and G with a kind of geometric distribution given by

$$\begin{aligned} \mathbb{P}(G = 2k) &= \left(1 - \mathrm{e}^{-c_1}\right) \mathrm{e}^{-k(c_0 + c_1)}, \\ \mathbb{P}(G = 2k+1) &= \left(1 - \mathrm{e}^{-c_0}\right) \mathrm{e}^{-k(c_0 + c_1) - c_1}, \end{aligned}$$

for $k \in \mathbb{N}_0$ are all independent.

Changing the definition of r_k for even and odd $k \in \mathbb{N}_0$ yields an analogous formula for L_+, the Laplace transform of τ_n conditional to $\tau_{n-1} = s$ and $(-1)^{\lfloor s \rfloor + n} = 1$, and all three cases below. Pay attention to the changed order of c_0 and c_1 in the third case.

Remark 3.4. The last case requires further explanations. Since the limit of $r_k T$ is finite and not equal to zero for $k = 0, 1$, the time scale T was chosen correctly to observe all jumps regardless of the original well. Suppose $(-1)^{\lfloor s \rfloor + n} = -1$ then the moving particle lies within the bigger well. The state -1 represents this well until $t = T(\lfloor s \rfloor + 1)$. Small values of E_1 indicate the possibility of a jump before $t = T(\lfloor s \rfloor + 1)$. The exact jump time is governed by the independent random variable $E_2 \sim \operatorname{Exp}([s, \lfloor s \rfloor + 1), c_0)$. If E_1 is bigger than $\lfloor s \rfloor + 1 - s$, the jump interval will be selected by the random variable G. Even values represent the exit from the smaller well and odd values indicate that the approximated process lies within the bigger well before the n-th jump occurs. For the concrete jump times again exponentially distributed random variables X and Y are exploited.

We owe the proof of Lemma 3.3.

Proof. Insert the relevant density f_- computed in Lemma 3.2 to prove the formula of $L_-(x)$

$$\begin{aligned} L_-(x) =& \mathbb{E}\left(\mathrm{e}^{-x\tau_n} \middle| \tau_{n-1} = s, (-1)^{\lfloor s \rfloor + n} = -1\right) \\ =& \frac{r_0 T}{r_0 T + x}\left(-\mathrm{e}^{-xt - r_0 T(t-s)}\Big|_s^{\lfloor s \rfloor + 1}\right) \\ &+ \mathrm{e}^{-r_0 T(\lfloor s \rfloor + 1 - s)} \sum_{k=0}^{\infty} \frac{r_{k+1} T}{r_{k+1} T + x}\, \mathrm{e}^{-\sum_{i=1}^{k} r_i T} \left(-\mathrm{e}^{-xt - r_{k+1} T(t - \lfloor s \rfloor - k - 1)}\Big|_{\lfloor s \rfloor + k + 1}^{\lfloor s \rfloor + k + 2}\right). \end{aligned} \tag{3.3}$$

If $r_k T$ for $k = 0, 1$ tends to zero as ε converges to zero, then for $x > 0$ the Laplace transform $L_-(x)$ also converge to zero due to the dominated convergence theorem of Lebesgue and the estimate

$$\sum_{k=0}^{\infty} \frac{r_{k+1} T}{r_{k+1} T + x}\, \mathrm{e}^{-\sum_{i=1}^{k} r_i T - x(\lfloor s \rfloor + k + 1)} \left(1 - \mathrm{e}^{-r_{k+1} T - x}\right) \leqslant \sum_{k=0}^{\infty} \mathrm{e}^{-xk} = \frac{1}{1 - \mathrm{e}^{-x}} < \infty.$$

But $L \equiv 0$ is the unique Laplace transform of the null measure.

In the case where T diverges much faster to infinity than r_k converges to zero for $k = 0, 1$, the first summand of $L_-(x)$ which converges to e^{-xs}, is decisive. Again because of the dominating convergence and the convergence of $\mathrm{e}^{-r_k T}$ to zero for $k = 0, 1$, the series in the last line of formula (3.3) vanishes in the small noise limit. Thus the limit law of τ_n is the Dirac measure in the starting time s.

For the last case further calculations are necessary:

$$\begin{aligned} L_-(x) =& \frac{r_0 T}{r_0 T + x}\left(\mathrm{e}^{-xs} - \mathrm{e}^{-x(\lfloor s \rfloor + 1) - r_0 T(\lfloor s \rfloor + 1 - s)}\right) \\ &+ \mathrm{e}^{-r_0 T(\lfloor s \rfloor + 1 - s)} \frac{r_1 T}{r_1 T + x}\, \mathrm{e}^{-x(\lfloor s \rfloor + 1)} \sum_{k=0}^{\infty} \mathrm{e}^{-x2k} \left(1 - \mathrm{e}^{-r_1 T - x}\right) \mathrm{e}^{-kT(r_1 + r_0)} \\ &+ \mathrm{e}^{-r_0 T(\lfloor s \rfloor + 1 - s)} \frac{r_0 T}{r_0 T + x}\, \mathrm{e}^{-x(\lfloor s \rfloor + 1) - x} \sum_{k=0}^{\infty} \mathrm{e}^{-x2k} \left(1 - \mathrm{e}^{-r_0 T - x}\right) \mathrm{e}^{-kT(r_1 + r_0) - r_1 T}. \end{aligned}$$

Thus the following is true

$$\begin{aligned} \lim_{\varepsilon \to 0} L_-(x) =& \frac{c_0}{c_0 + x}\left(\mathrm{e}^{-xs} - \mathrm{e}^{-x(\lfloor s \rfloor + 1) - c_0(\lfloor s \rfloor + 1 - s)}\right) \\ &+ \mathrm{e}^{-c_0(\lfloor s \rfloor + 1 - s) - x(\lfloor s \rfloor + 1)} \frac{c_1}{c_1 + x} \frac{1 - \mathrm{e}^{-c_1 - x}}{1 - \mathrm{e}^{-c_1 - c_0 - 2x}} \\ &+ \mathrm{e}^{-c_0(\lfloor s \rfloor + 1 - s)} \frac{c_0}{c_0 + x}\, \mathrm{e}^{-x(\lfloor s \rfloor + 1) - x} \frac{1 - \mathrm{e}^{-c_0 - x}}{1 - \mathrm{e}^{-c_1 - c_0 - 2x}}. \end{aligned}$$

Use $E_1 \sim \mathrm{Exp}(c_0)$ to compute the Laplace transform of $E_2 \sim \mathrm{Exp}([s, \lfloor s \rfloor + 1), c_0)$

$$\begin{aligned}
\mathbb{E}\,\mathrm{e}^{-xE_2} &= \int_s^{\lfloor s \rfloor+1} \mathrm{e}^{-xt} \frac{c_0\,\mathrm{e}^{-c_0 t}}{\mathrm{e}^{-as} - \mathrm{e}^{-c_0(\lfloor s \rfloor+1)}}\,\mathrm{d}t \\
&= \frac{\mathrm{e}^{c_0 s}}{1-\mathrm{e}^{-c_0(\lfloor s \rfloor+1-s)}} \frac{c_0}{c_0+x} \left(-\mathrm{e}^{-(c_0+x)t} \Big|_s^{\lfloor s \rfloor+1} \right) \\
&= \frac{c_0}{c_0+x} \frac{\mathrm{e}^{-xs} - \mathrm{e}^{-c_0(\lfloor s \rfloor+1-s)-x(\lfloor s \rfloor+1)}}{\mathbb{P}(E_1 \leqslant \lfloor s \rfloor + 1 - s)}.
\end{aligned}$$

This immediately justifies the first summand of $\lim_{\varepsilon\to 0} L_-(x)$. Due to the independence of $X \sim \mathrm{Exp}([G, G+1), c_1)$ and E_1 it holds

$$\begin{aligned}
&\mathbb{E}\left(\mathrm{e}^{-x(X+\lfloor s \rfloor+1)}\, \mathbf{1}_{\{G \text{ is even}\}}\, \mathbf{1}_{\{E_1 \geqslant \lfloor s \rfloor+1-s\}} \right) \\
&= \mathrm{e}^{-c_0(\lfloor s \rfloor+1-s)-x(\lfloor s \rfloor+1)} \sum_{k=0}^{\infty} \left(1-\mathrm{e}^{-c_1}\right) \mathrm{e}^{-k(c_0+c_1)} \int_{2k}^{2k+1} \mathrm{e}^{-xt} \frac{c_1\,\mathrm{e}^{-c_1 t}}{\mathrm{e}^{-c_1 2k} - \mathrm{e}^{-c_1(2k+1)}}\,\mathrm{d}t \\
&= \mathrm{e}^{-c_0(\lfloor s \rfloor+1-s)-x(\lfloor s \rfloor+1)} \left(1-\mathrm{e}^{-c_1}\right) \sum_{k=0}^{\infty} \frac{c_1}{c_1+x} \frac{\mathrm{e}^{-(x+c_1)2k} - \mathrm{e}^{-(x+c_1)(2k+1)}}{\mathrm{e}^{-c_1 2k}(1-\mathrm{e}^{-b})} \mathrm{e}^{-k(a+b)} \\
&= \mathrm{e}^{-c_0(\lfloor s \rfloor+1-s)-x(\lfloor s \rfloor+1)} \frac{c_1}{c_1+x} \frac{1-\mathrm{e}^{-c_1-x}}{1-\mathrm{e}^{-c_0-c_1-2x}},
\end{aligned}$$

which yields the second summand of the limit of $L_-(x)$. The third summand follows from analogous observations and the use of $Y \sim \mathrm{Exp}([G, G+1), c_0)$ and $\{G \text{ is odd}\}$ instead of X and $\{G \text{ is even}\}$. □

The random signal K with conditional jump time densities $f_\pm$ looks quite periodic if the most mass of the density f_- is concentrated around odd multiples of T and if f_+ has large values around multiples of $2T$. A natural quantity of optimal tuning could be the function

$$\mathbb{P}_{-1}(\tau_1 \in [1-\delta, 1+\delta]),$$

for $\delta > 0$ small and fixed and initial value -1. In [19] Herrmann and Imkeller considered the small parameter ε as fixed and searched for the correct time scale to maximize this quantity. In our scenario the periodicity $2T$ is given and the behaviour in the small noise limit is of interest. If a quality measure of tuning depending on T and ε admits an optimal time scale $T_0 = T_0(\varepsilon)$ and an optimal noise intensity $\varepsilon_0 = \varepsilon_0(T)$, this a remarkable outcome which should not be taken for granted. This phenomenon was given the special name double stochastic resonance. In the sequel we will paraphrase how to define T_0 and ε_0 and even generalize these definitions for the probability to jump within the interval $[2k+1-\delta, 2k+1+\delta]$ for $k \in \mathbb{N}_0$, because a missed jump in the normalized time interval $[1-\delta, 1+\delta]$ should be made up near odd integers to justify a periodic appearance.

Proposition 3.5. *Fix $\delta \in (0,1)$ small and assume $k \in \mathbb{N}_0$, $T > 0$ and $\varepsilon > 0$. Then it holds:*

$$m(T,\varepsilon) := \mathbb{P}_{-1}(\tau_1 \in [2k+1-\delta, 2k+1+\delta]) = \mathrm{e}^{-\varphi_K T(k+1)-\psi_K T k} \left(\mathrm{e}^{\varphi_K T\delta} - \mathrm{e}^{-\psi_K T\delta} \right).$$

For any $\varepsilon > 0$ the function $T \mapsto m(T,\varepsilon)$ attains its maximum at

$$T_{k,\delta}(\varepsilon) = \frac{1}{\varepsilon^\alpha l(\varepsilon^{-1})\delta(a^{-\alpha}+b^{-\alpha})} \log \frac{(k+1)a^{-\alpha}+(k+\delta)b^{-\alpha}}{(k+1-\delta)a^{-\alpha}+kb^{-\alpha}}$$

with

$$T_k(\varepsilon) := \lim_{\delta\to 0} T_{k,\delta}(\varepsilon) = \frac{1}{\varepsilon^\alpha l(\varepsilon^{-1})} \frac{1}{(k+1)a^{-\alpha}+kb^{-\alpha}}.$$

Assume $h(\varepsilon) = \varepsilon^\alpha l(\varepsilon^{-1})$ is continuous and strictly monotone for small ε, so that its inverse h^{-1} exists. Then for any $T > 0$ the map $\varepsilon \mapsto m(T,\varepsilon)$ attains its maximum at

$$\varepsilon_{k,\delta}(T) = h^{-1}\left(\frac{1}{T\delta(a^{-\alpha}+b^{-\alpha})} \log \frac{(k+1)a^{-\alpha}+kb^{-\alpha}+\delta\, b^{-\alpha}}{(k+1)a^{-\alpha}+kb^{-\alpha}-\delta\, a^{-\alpha}}\right)$$

and as $\delta \to 0$ the position of the maximum $\varepsilon_{k,\delta}(T)$ tends to $\varepsilon_k(T) := h^{-1}\left(\frac{1}{T((k+1)a^{-\alpha}+kb^{-\alpha})}\right)$. In particular, with $l \equiv 1$ one obtains the following in the limit $\delta \to 0$

$$\varepsilon_{k,\delta}(T) = \left(\frac{1}{T\delta(a^{-\alpha}+b^{-\alpha})} \log \frac{(k+1)a^{-\alpha}+kb^{-\alpha}+\delta\, b^{-\alpha}}{(k+1)a^{-\alpha}+kb^{-\alpha}-\delta\, a^{-\alpha}}\right)^{1/\alpha} \to \frac{1}{(T((k+1)a^{-\alpha}+kb^{-\alpha}))^{1/\alpha}}.$$

This result confirms the theoretical suggestions that the optimal periodicity $2T$ is equal to $\frac{c}{\varepsilon^\alpha l(\varepsilon^{-1})}$ for a suitable $c > 0$. Figure 3.4 and 3.5 serve as illustration.

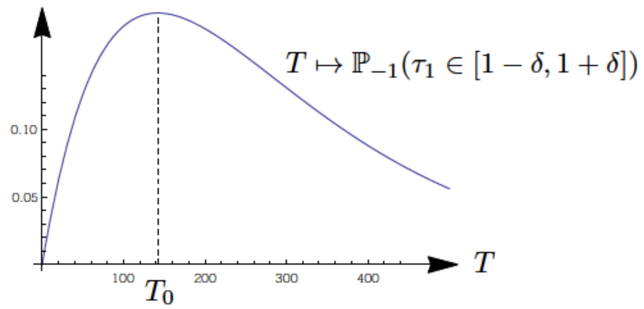


Figure 3.4: For $l \equiv 1$, $a = 5$, $b = 1$, $\alpha = 1$, $\delta = 0.1$ and $\varepsilon = 0.03$ the function $T \mapsto \mathbb{P}_{-1}(\tau_1 \in [1-\delta, 1+\delta])$ has a maximum at $T_0 = 141.896$.

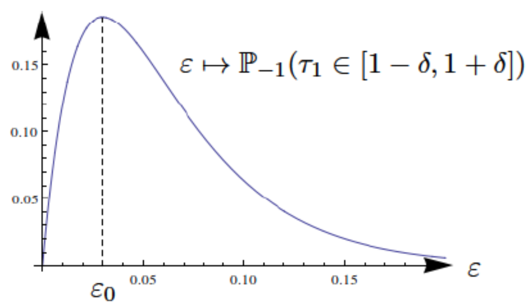


Figure 3.5: For $l \equiv 1$, $a = 5$, $b = 1$, $\alpha = 1$, $\delta = 0.1$ and $T = 142$ the function $\varepsilon \mapsto \mathbb{P}_{-1}(\tau_1 \in [1-\delta, 1+\delta])$ has a maximum at $\varepsilon_0 = 0.029978$.

Proof. Apply the formula for the conditional density given in Lemma (3.2) to get

$$\begin{aligned}
m(\varepsilon, T) &= \mathrm{e}^{-kT(\varphi_K+\psi_K)} \int_{2k+1-\delta}^{2k+1} \varphi_K \, T \, \mathrm{e}^{-\varphi_K T(t-2k)} \, \mathrm{d}t \\
&\quad + \mathrm{e}^{-kT(\varphi_K+\psi_K)-T\varphi_K} \int_{2k+1}^{2k+1+\delta} \psi_K \, T \, \mathrm{e}^{-\psi_K T(t-2k-1)} \, \mathrm{d}t \\
&= \mathrm{e}^{-kT(\varphi_K+\psi_K)} \left[\mathrm{e}^{-T\varphi_K(1-\delta)} - \mathrm{e}^{-\varphi_K T} + \mathrm{e}^{-\varphi_K T} - \mathrm{e}^{-\psi_K T\delta - \varphi_K T} \right] \\
&= -\mathrm{e}^{-\varphi_K T - kT(\varphi_K+\psi_K)} (\mathrm{e}^{\varphi_K T\delta} - \mathrm{e}^{-\psi_K T\delta}).
\end{aligned}$$

Obviously it holds $m(\varepsilon, 0) = 0$, $\lim_{T\to\infty} m(\varepsilon, T) = 0$ and $m(\varepsilon, T) \geqslant 0$ for $T \geqslant 0$. Therefore a single extremum must be a maximum. It follows

$$\begin{aligned}
\frac{\partial}{\partial T} m(\varepsilon, T) = &- ((k+1-\delta)\varphi_K - k\psi_K) \, \mathrm{e}^{-T((k+1-\delta)\varphi_K + k\psi_K)} \\
&+ ((k+1)\varphi_K + (k+\delta)\psi_K) \, \mathrm{e}^{-T((k+1)\varphi_K + (k+\delta)\psi_K)}
\end{aligned}$$

and from $\frac{\partial}{\partial T} m(\varepsilon, T) = 0$ we deduce

$$\frac{(k+1)\varphi_K + (k+\delta)\psi_K}{(k+1-\delta)\varphi_K + k\psi_K} = \mathrm{e}^{-T((k+1)\varphi_K + (k+\delta)\psi_K) + T((k+1-\delta)\varphi_K + k\psi_K)} = \mathrm{e}^{T\delta(\varphi_K+\psi_K)}.$$

Thus $T \mapsto m(\varepsilon, T)$ attains its largest value at $T_{k,\delta}$ as given in the proposition. If δ tends to zero, $T_{k,\delta}$ converges to the inverse of $((k+1)a^{-\alpha} + kb^{-\alpha})\varepsilon^\alpha l(\varepsilon^{-1})$, because l'Hôspital's rule implies

$$\begin{aligned}
&\lim_{\delta\to 0} \frac{1}{\delta(a^{-\alpha}+b^{-\alpha})} \log \frac{(k+1)a^{-\alpha} + (k+\delta)\, b^{-\alpha}}{(k+1-\delta)a^{-\alpha} + kb^{-\alpha}} \\
&= \lim_{\delta\to 0} \frac{\left[(k+1-\delta)a^{-\alpha} + kb^{-\alpha}\right] \left[(ab)^{-\alpha}(k+1-\delta) + kb^{-2\alpha} + a^{-2\alpha}(k+1) + (ab)^{-\alpha}(k+\delta)\right]}{(a^{-\alpha}+b^{-\alpha}) \left[(k+1)a^{-\alpha} + (k+\delta)b^{-\alpha}\right] \left[(k+1-\delta)a^{-\alpha} + kb^{-\alpha}\right]^2} \\
&= \frac{1}{(k+1)a^{-\alpha} + kb^{-\alpha}}.
\end{aligned}$$

The same arguments as in the first part of the proof and the monotonicity of h yield a single maximum of $\varepsilon \mapsto m(\varepsilon, T)$ at $\varepsilon_{k,\delta}$ given above. □

3.1.2 Continuous infinitesimal generator

In the last subsection the lack of continuity unnecessarily complicates all calculations and does not depict the movement of a noisy particle in a periodically varying potential properly. A more natural way to approximate the solution of the stochastic differential equation (3.1) with periodic drift is an inhomogeneous Markov chain $C = (C_t^\varepsilon)_{t\geqslant 0}$ on the state space $\{-1, 1\}$ with infinitesimal generator given by

$$Q^C(t) = \begin{pmatrix} -\varphi_C(t) & \varphi_C(t) \\ \varphi_C(t+T) & -\varphi_C(t+T) \end{pmatrix}, \; t \geqslant 0,$$

where

$$\varphi_C(t) = \frac{\varepsilon^\alpha l(\varepsilon^{-1})}{p(t/2T)^\alpha}$$

for $t \geqslant 0$ and a continuous, 1-periodic function p with $p(0) = \max_{t\geqslant 0} p(t)$ and $p\left(\frac{1}{2}\right) = \min_{t\geqslant 0} p(t) > 0$. The time lag T and the periodicity of p are chosen to realize the switching of the roles of the states -1 and 1 after a half of a period. For example, we can use

$$p(t) = \frac{a+b}{2} + \frac{a-b}{2}\cos 2\pi t \tag{3.4}$$

that continuously varies between the values a and $b < a$. Then at the beginning, -1 can be identified with a bigger well. For t tending to T the well gets smaller and smaller, until it reaches its smallest width at $t = T$. After that it widens again. At $t = 2T$ the original situation is re-established.

Definition 3.6. *Let $T_0 = 0$ and T_n, $n \in \mathbb{N}$ be the n-th jump time of C from -1 to 1 or the other way around and $\tau_n = \frac{T_n}{T}$, $n \in \mathbb{N}_0$.*

Again we are interested in the conditional density of τ_n given the Markov chain lies within the smaller or the bigger well. Justified by the continuity of Q^C the law of τ_n is uniquely determined by $\tau_{n-1} = s$ and the value of $(-1)^n$. While $(-1)^n = -1$ demands the use of transition rate φ_C, $(-1)^n = 1$ indicates the transition rate equals to $\varphi_C(T + \cdot)$. The form of the conditional density $f_\pm$ of τ_n still shows peaks at even respectively odd integers, but now it is continuous (Figure 3.6).

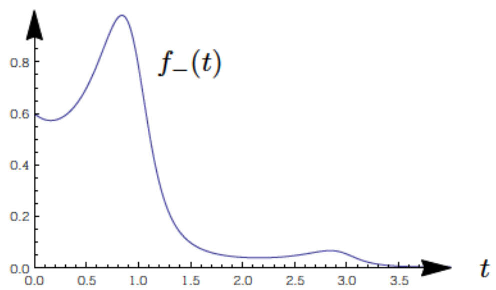


Figure 3.6: The density of τ_1 given $C_0 = -1$ with $l \equiv 1$, $a = 5$, $b = 1$, $\alpha = 1$, p as in formula (3.4) and $T = \frac{(a+b)/2}{\varepsilon}$.

Lemma 3.7. *The n-th normalized jump time τ_n of C possesses the following density given that $\tau_{n-1} = s$ and $(-1)^n = -1$*

$$f_-(t) = \mathbf{1}_{[s,\infty)}(t)\, T\varphi_C(Tt)\, \mathrm{e}^{-T\int_s^t \varphi_C(Tr)\,\mathrm{d}r}, \quad t \geqslant s$$

and

$$f_+(t) = \mathbf{1}_{[s,\infty)}(t)\, T\varphi_C(T(t+1))\, \mathrm{e}^{-T\int_s^t \varphi_C(T(r+1))\,\mathrm{d}r}, \quad t \geqslant s,$$

if $\tau_{n-1} = s$ and $(-1)^n = 1$.

Analogously to the previous subsection an analysis of the asymptotic behaviour of C through computation of the conditional Laplace transform of τ_n is possible. The following lemma partly follows the lines of [19] (Lemma 3 in Chapter 2).

Lemma 3.8. *The conditional Laplace transform of τ_n given $\tau_{n-1} = s$ and $(-1)^n = -1$ is*

$$L_-(x) = \frac{\int_s^{s+2} T\varphi_C(Tt)\,\mathrm{e}^{-xt - T\int_s^t \varphi_C(Tr)\,\mathrm{d}r}\mathrm{d}t}{1 - \mathrm{e}^{-2x - T\int_0^2 \varphi_C(Tr)\,\mathrm{d}r}}, \quad x \geqslant 0.$$

(i) If $\lim_{\varepsilon\to 0} T\varphi_C(Tt) = 0$ uniformly for all $t \in [0,2]$, the weak limit of the conditional law of τ_n is the null measure.

(ii) If $\lim_{\varepsilon\to 0} T\varphi_C(Tt) = \infty$ uniformly for all $t \in [0,2]$, the conditional law of τ_n weakly tends to the Dirac measure in s.

(iii) Assume $T = \frac{c}{\varepsilon^\alpha l(\varepsilon^{-1})}$ for $c > 0$. Then τ_n behaves like a random variable X with density

$$f(t) = c\,p\left(\frac{t}{2}\right)^{-\alpha} \mathrm{e}^{-\int_s^t c\,p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}\mathbf{1}_{[s,\infty)}(t), \quad t \geqslant s.$$

A similar formula for L_+, the Laplace transform of τ_n conditional to $\tau_{n-1} = s$ and $(-1)^n = 1$, and the same behaviour in the different cases follow by replacing $\varphi_C(T\cdot)$ with $\varphi_C(T(\cdot+1))$ respectively $p(\cdot/2)$ with $p((\cdot+1)/2)$.

Proof. From the 2-periodicity of $\varphi_C(T\cdot)$ we derive

$$\begin{aligned}
L_-(x) &= \mathbb{E}\left(\mathrm{e}^{-x\tau_n}\middle|\tau_{n-1} = s, (-1)^n = -1\right)\\
&= \sum_{k=0}^{\infty}\int_{s+2k}^{s+2(k+1)} \mathrm{e}^{-xt}\,T\varphi_C(Tt)\,\mathrm{e}^{-T\int_s^{s+2k}\varphi_C(Tr)\,\mathrm{d}r - T\int_{s+2k}^t \varphi_C(Tr)\,\mathrm{d}r}\mathrm{d}t\\
&= \sum_{k=0}^{\infty}\mathrm{e}^{-Tk\int_0^2\varphi_C(Tr)\,\mathrm{d}r - x2k}\int_s^{s+2} T\varphi_C(Tt)\mathrm{e}^{-xt - T\int_s^t\varphi_C(Tr)\,\mathrm{d}r}\mathrm{d}t
\end{aligned}$$

and basic knowledge about geometric series verifies the formula of $L_-(x)$.
Assume $T\varphi_C(Tt)$ uniformly in t tends to zero as ε converges to zero. Thus $T\varphi_C(Tt)$ is bounded by a small constant δ for all small noise intensities and all $t \geqslant 0$. Since additionally the formula of $L_-(x)$ contains only integrals over finite intervals, the denominator of $L_-(x)$ admits $1 - \mathrm{e}^{-2x}$ as limit and the numerator tends to zero as ε converges to zero. This yields the null measure as the law in the small noise limit.
If $T\varphi_C(Tt)$ diverges uniformly for at least t taken from a small interval, the denominator of $L_-(x)$ converges to 1, because the Lemma of Fatou guarantees

$$\liminf_{\varepsilon\to 0}\int_s^{s+2} T\varphi_C(Tt)\,\mathrm{d}t \geqslant \int_s^{s+2}\liminf_{\varepsilon\to 0} T\varphi_C(Tt)\,\mathrm{d}t = \infty.$$

Because the convergence of $T\varphi_C(Tt)$ is uniform in all $t \in [0,2]$, the following holds

$$\frac{T\varphi_C(Tt)}{T\varphi_C(Tt)+x} = \frac{1}{1+x(T\varphi_C(Tt))^{-1}} \geqslant 1 - \mathrm{o}_\varepsilon(1)$$

for all $t \geqslant 0$, while $\mathrm{o}_\varepsilon(1)$ tends to zero in the small noise limit. Through this lower bound, exact computation and another application of the result of Fatou we obtain

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\int_s^{s+2} T\varphi_C(Tt)\,\mathrm{e}^{-xt-T\int_s^t \varphi_C(Tr)\,\mathrm{d}r}\mathrm{d}t \\
&\geqslant \lim_{\varepsilon\to 0}(1-\mathrm{o}_\varepsilon(1))\int_s^{s+2}(T\varphi_C(Tt)+x)\,\mathrm{e}^{-xt-T\int_s^t \varphi_C(Tr)\,\mathrm{d}r}\mathrm{d}t \\
&= \lim_{\varepsilon\to 0}\mathrm{e}^{-xs} - \mathrm{e}^{-x(s+2)-\int_s^{s+2} T\varphi_C(Tr)\,\mathrm{d}r} \\
&= \mathrm{e}^{-xs}.
\end{aligned}$$

Because $\frac{T\varphi_C(Tt)}{T\varphi_C(Tt)+x}$ is smaller or equal to 1, e^{-xs} is also an upper bound for the limit of the numerator. This proves the limit e^{-xs} for $L_-(x)$ in the second case. This Laplace transform corresponds to the Dirac measure in s.

To calculate the Laplace transform of the random variable X given by the density f in (iii) it suffices to exploit the periodicity of $c\,p\left(\frac{t}{2}\right)^{-\alpha}$ again to get

$$\mathbb{E}\,\mathrm{e}^{-xX} = \frac{\int_s^{s+2} c\,p\left(\frac{t}{2}\right)^{-\alpha}\mathrm{e}^{-xt-\int_s^t c\,p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}\mathrm{d}t}{1-\mathrm{e}^{-2x-\int_0^2 c\,p\left(\frac{t}{2}\right)^{-\alpha}\mathrm{d}t}},$$

which does not differ from $L_-(x)$ in the third case. □

The interpretation of this result is similar to the case with piecewise constant infinitesimal generator. The smallness of a time scale T of order $\varepsilon^{-\beta}$ with $\beta < \alpha$ prevents recording any jump, but if $\beta > \alpha$ the unit of length T goes by too quickly and therefore only an instantaneous jump is observed. For $T = \frac{c}{\varepsilon^\alpha l(\varepsilon^{-1})}$ the Laplace transform is independent of ε. This supports the importance of the time scale $\varepsilon^{-\alpha}$. For an explicit calculation of the optimal time scale we copy the procedure of the last subsection.

Proposition 3.9. *Fix $\delta \in (0,1)$ small and assume $k \in \mathbb{N}_0$, $\varepsilon > 0$ and $T > 0$. The function $T \mapsto \mathbb{P}_{-1}(\tau_1 \in [2k+1-\delta, 2k+1+\delta])$ attains its largest value at*

$$T_{k,\delta}(\varepsilon) = \frac{1}{\varepsilon^\alpha l(\varepsilon^{-1})\int_{2k+1-\delta}^{2k+1+\delta} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}\log\frac{\int_0^{2k+1+\delta} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}{\int_0^{2k+1-\delta} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}.$$

As δ tends to zero the position of the maximum $T_{k,\delta}(\varepsilon)$ converges to $T_k(\varepsilon) := \left(\varepsilon^\alpha l(\varepsilon^{-1})\int_0^{2k+1} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r\right)^{-1}$. Assume $h(\varepsilon) = \varepsilon^\alpha l(\varepsilon^{-1})$ is continuous and strictly monotone for small ε. Hence the inverse h^{-1} exists for small ε. Then $\varepsilon \mapsto \mathbb{P}_{-1}(\tau_1 \in [2k+1-\delta, 2k+1+\delta])$ attains its largest value at

$$\varepsilon_{k,\delta}(T) = h^{-1}\left(\frac{1}{T\int_{2k+1-\delta}^{2k+1+\delta} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}\log\frac{\int_0^{2k+1+\delta} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}{\int_0^{2k+1-\delta} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}\right).$$

As $\delta \to 0$ *the maximum position* $\varepsilon_{k,\delta}(T)$ *tends to* $\varepsilon_k(T) := h^{-1}\left(\left[T\int_0^{2k+1} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r\right]^{-1}\right)$. *In particular, with* $l \equiv 1$ *this implies*

$$\varepsilon_{k,\delta}(T) = \left(\frac{1}{T\int_{2k+1-\delta}^{2k+1+\delta} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}\log\frac{\int_0^{2k+1+\delta} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}{\int_0^{2k+1-\delta} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r}\right)^{1/\alpha} \to \frac{1}{\left(T\int_0^{2k+1} p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r\right)^{1/\alpha}}$$

as $\delta \to 0$.

Proof. Insert the density of Lemma 3.7 and define $I(t) = \int_0^t p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r$ to obtain

$$\begin{aligned}\mathbb{P}_{-1}(\tau_1 \in [2k+1-\delta, 2k+1+\delta]) &= \int\limits_{2k+1-\delta}^{2k+1+\delta} T\varepsilon^\alpha l(\varepsilon^{-1})\,\mathrm{e}^{-T\varepsilon^\alpha l(\varepsilon^{-1})I(t)}\,\mathrm{d}t \\ &= \mathrm{e}^{-T\varepsilon^\alpha l(\varepsilon^{-1})I(2k+1-\delta)} - \mathrm{e}^{-T\varepsilon^\alpha l(\varepsilon^{-1})I(2k+1+\delta)}.\end{aligned} \tag{3.5}$$

Differentiating with respect to T and equating the deviation with zero yields the formula of $T_{k,\delta}$. It suffices to revive arguments of the proof of Lemma 3.5 to deduce there is a maximum. The smoothness of p justifies $I(2k+1+\delta) \to I(2k+1)$ as $\delta \to 0$. The rule of l'Hospital implies

$$\begin{aligned}&\lim_{\delta\to 0}\frac{1}{I(2k+1+\delta) - I(2k+1-\delta)}\log\frac{I(2k+1+\delta)}{I(2k+1-\delta)} \\ &= \lim_{\delta\to 0}\frac{1}{(2k+1)\left(p\left(\frac{1+\delta}{2}\right)^{-\alpha} + p\left(\frac{1-\delta}{2}\right)^{-\alpha}\right)}\frac{I(2k+1-\delta)}{I(2k+1+\delta)} \\ &\qquad\times\frac{(2k+1)p\left(\frac{1+\delta}{2}\right)^{-\alpha}I(2k+1-\delta) + (2k+1)p\left(\frac{1-\delta}{2}\right)^{-\alpha}I(2k+1+\delta)}{I(2k+1-\delta)^2}.\end{aligned}$$

This expression converges to $(I(2k+1))^{-1}$. The proof of the remaining part of this proposition does not reveal new techniques, hence it is omitted. □

Remark 3.10. Again the phenomenon of double stochastic resonance is observed (for example see Figures 3.7 and 3.8) which underpins the application of the used quality measure of tuning. In Subsection 3.4.2 a similar measure is analysed for the jump diffusions X^ε and Z^ε solving equations (3.1) and (3.2). The same optimal values $T(\varepsilon)$ for the half period are verified.

To get an impression of the behaviour of $\mathbb{P}_{-1}(\tau_1 \in [1-\delta, 1+\delta])$ in terms of δ, use the definition $I(t) = \int_0^t p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r$, apply the series expansion and the mean value theorem to justify the succeeding approximations for small δ

$$\begin{aligned}\mathbb{P}_{-1}(\tau_1 \in [1-\delta, 1+\delta]) &= \mathrm{e}^{-T\varepsilon^\alpha l(\varepsilon^{-1})I(1-\delta)} - \mathrm{e}^{-T\varepsilon^\alpha l(\varepsilon^{-1})I(1+\delta)} \\ &= \mathrm{e}^{-T\varepsilon^\alpha l(\varepsilon^{-1})(I(1)-\delta I'(1)+\mathrm{O}(\delta^2))} - \mathrm{e}^{-T\varepsilon^\alpha l(\varepsilon^{-1})(I(1)+\delta I'(1)+\mathrm{O}(\delta^2))} \\ &= \mathrm{e}^{-T\varepsilon^\alpha l(\varepsilon^{-1})I(1)}(\mathrm{e}^{T\varepsilon^\alpha l(\varepsilon^{-1})I'(1)\delta} - \mathrm{e}^{-T\varepsilon^\alpha l(\varepsilon^{-1})I'(1)\delta}) + T\varepsilon^\alpha l(\varepsilon^{-1})\mathrm{O}(\delta^2) \\ &= 2\delta\, T\varepsilon^\alpha l(\varepsilon^{-1})\, p\left(\frac{1}{2}\right)^{-\alpha}\mathrm{e}^{-T\varepsilon^\alpha l(\varepsilon^{-1})\int_0^1 p\left(\frac{r}{2}\right)^{-\alpha}\mathrm{d}r} + T\varepsilon^\alpha l(\varepsilon^{-1})\mathrm{O}(\delta^2).\end{aligned}$$

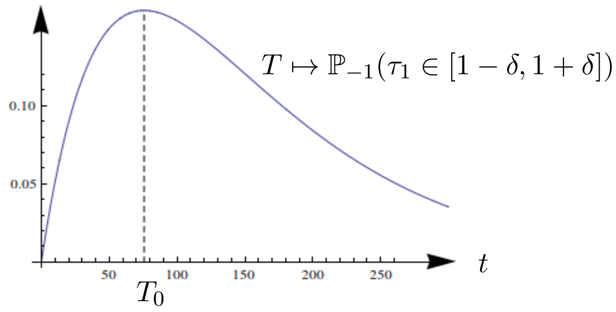


Figure 3.7: For $l \equiv 1$, $a = 5$, $b = 1$, $\alpha = 1$, $\delta = 0.1$ and $\varepsilon = 0.03$ the function $T \mapsto \mathbb{P}_{-1}(\tau_1 \in [1-\delta, 1+\delta])$ has a maximum at $T_0 = 75.7362$.

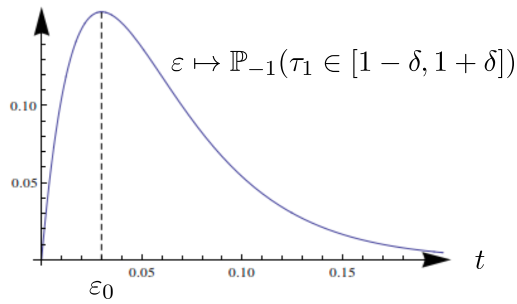


Figure 3.8: For $l \equiv 1$, $a = 5$, $b = 1$, $\alpha = 1$, $\delta = 0.1$ and $T = 76$ the function $\varepsilon \mapsto \mathbb{P}_{-1}(\tau_1 \in [1-\delta, 1+\delta])$ has a maximum at $\varepsilon_0 = 0.0298959$.

From $T = \left(\varepsilon^\alpha l(\varepsilon^{-1}) \int_0^1 p\left(\frac{r}{2}\right)^{-\alpha} \mathrm{d}r\right)^{-1}$ one derives

$$\mathbb{P}_{-1}(\tau_1 \in [1-\delta, 1+\delta]) = \delta \frac{p\left(\frac{1}{2}\right)^{-\alpha}}{\mathrm{e} \int_0^{1/2} p\,(r)^{-\alpha}\,\mathrm{d}r} + \mathrm{O}(\delta^2).$$

3.2 The jump diffusion and the decomposition of jumps

In the remaining part of this chapter we devote our attention to the jump diffusions solving equation (3.1) respectively (3.2).

Consult Section 2.1 to justify the existence of $2T$-periodic solutions $p_\pm$ in the vicinity of the varying minimum positions $m_\pm(t)$ of U and a period solution p_0 near the saddle at $m_0(t)$.

Definition 3.11. *Define the basin of attraction of the solution $x(t) = x(t; x_0, t_0)$ of $\dot{x}(t) = -\nabla U\left(x(t), \frac{t}{2T}\right)$, $t \geqslant t_0$ with initial value $x(t_0) = x_0$ by*

$$\Omega_\pm(t_0) = \left\{x_0 \in \mathbb{R}^d \colon \|x(t; x_0, t_0) - p_\pm(t)\| \to 0 \text{ as } t \to \infty\right\}.$$

and $\Gamma(t_0) = \mathbb{R}^d \backslash (\Omega_-(t_0) \cup \Omega_+(t_0))$ is the separating manifold called separatrix.

The sets $\Omega_\pm(t_0)$ and $\Gamma(t_0)$ are invariant and $2T$-periodic. The basins of attraction are open sets. For $d = 1$ the field which determines the solution $\hat{x}$ of $\frac{1}{2T}\frac{\mathrm{d}}{\mathrm{d}s}\hat{x}(s) = -\nabla U(\hat{x}(s), s)$

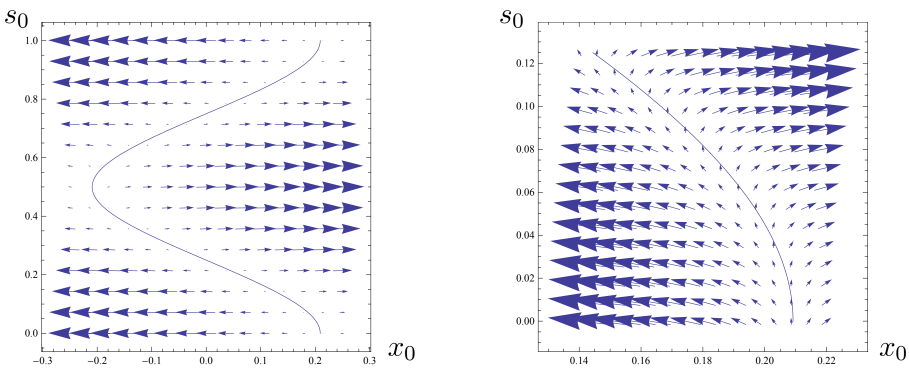


Figure 3.9: The gradient field of $\frac{1}{100}\frac{\mathrm{d}}{\mathrm{d}s}\hat{x}(s) = -\nabla U(\hat{x}(s), s)$, $s \geqslant s_0$, $\hat{x}(s_0) = x_0$ with $\nabla U(x, s) = x^3 - x + 0.2\cos 2\pi s$ for different x_0 and s_0 is illustrated relatively to $m_0(\cdot)$ (solid line).

depending on different initial values is illustrated in Figure 3.9. Due to the periodic changing of the potential U there exist solutions that start left of the maximum but in fact tend to the periodic solution near the right minimum (and vice versa). This can be seen especially in the right picture of Figure 3.9 when an arrow starting left of the solid line modelling $m_0(s_0)$ points to the right.

Definition 3.12. *For $\gamma, \varepsilon > 0$ and $t_0 \geqslant 0$ define the reduced domains of attraction*

$$\Omega_\pm^{\varepsilon^\gamma}(t_0) = \left\{x_0 \in \mathbb{R}^d \colon B_{\varepsilon^\gamma}(x(t; x_0, t_0)) \subseteq \Omega_\pm(t),\ \textit{for all } t \geqslant t_0\right\},$$
$$\Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0) = \left\{x_0 \in \Omega_\pm^{\varepsilon^\gamma}(t_0) \colon B_{\varepsilon^{2\gamma}}(x(t; x_0, t_0)) \subseteq \Omega_\pm^{\varepsilon^\gamma}(t),\ \textit{for all } t \geqslant t_0\right\},$$
$$\Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t_0) = \left\{x_0 \in \Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0) \colon B_{\varepsilon^{2\gamma}}(x(t; x_0, t_0)) \subseteq \Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t),\ \textit{for all } t \geqslant t_0\right\},$$

and the corresponding enlarged separatrices

$$\Gamma^{\varepsilon^\gamma}(t_0) = \mathbb{R}^d \setminus \left(\Omega_-^{\varepsilon^\gamma}(t_0) \cup \Omega_+^{\varepsilon^\gamma}(t_0)\right),$$
$$\Gamma^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0) = \mathbb{R}^d \setminus \left(\Omega_-^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0) \cup \Omega_+^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0)\right),$$
$$\Gamma^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t_0) = \mathbb{R}^d \setminus \left(\Omega_-^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t_0) \cup \Omega_+^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t_0)\right).$$

The bounded basins of attraction and the truncated separatrix are given by

$$\Omega_{\pm,R}(t_0) = \Omega_\pm(t_0) \cap O_R, \quad \textit{and} \quad \Gamma_R(t_0) = \Gamma(t_0) \cap O_R,$$

for $R \geqslant R^* := R_U^*$ with R_U^* known from (U3). *Through intersection with an invariant set $O_R := O_R^U$, $R \geqslant R^*$, known from* (U3) *and the different reduced domains we obtain the bounded and reduced basins of attraction $\Omega_{\pm,R}^{\varepsilon^\gamma}(t_0), \Omega_{\pm,R}^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0)$ and $\Omega_{\pm,R}^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t_0)$. The definition of the bounded and enlarged separatrices $\Gamma_R^{\varepsilon^\gamma}(t_0)$, $\Gamma_R^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0)$ and $\Gamma_R^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t_0)$ is similar.*

Definition 3.13. *For $s \geqslant 0$ consider the homogeneous equation $\frac{\mathrm{d}}{\mathrm{d}t}\tilde{x}(t) = -\nabla U(\tilde{x}(t), s)$, $t \geqslant 0$. Let $\widetilde{\Omega}_\pm(s)$ be the basin of attraction of $m_\pm(s)$ corresponding to this equation.*

Finally, to guarantee a return time of the deterministic solution x that is logarithmic in ε and uniform in the initial value we have to exclude the neighbourhood of the separatrix and we rigorously define the importance of the time points kT, $k \in \mathbb{N}$, for the jump diffusions.

(Log) Assume for all $\gamma > 0$ there are $\varepsilon_0, c > 0$ such that for all $t_0 \geqslant 0$ and $x_0 \in \Omega_\pm^{\varepsilon^\gamma}(t_0)$ the solution $x(t; x_0, t_0)$ of (2.3) fulfills $\|x(t; x_0, t_0) - p_\pm(t)\| \leqslant \frac{1}{3}\varepsilon^{4\gamma}$ for all $t \geqslant c\gamma|\log \varepsilon| =: R_{\varepsilon^\gamma}$ and $\varepsilon < \varepsilon_0$.

(KT) Define the functions $J_\pm^I(t) = \mu\left(\left\{y \in \mathbb{R}^d : m_\pm(t) + g(m_\pm(t))y \in \widetilde{\Omega}_\mp(t)\right\}\right)$ and $J_\pm^M(t) = \mu\left(\left\{y \in \mathbb{R}^d : \varphi(1; m_\pm(t), y) \in \widetilde{\Omega}_\mp(t)\right\}\right)$. For $i = I, M$ assume $J_+^i(t)$ is maximal at $t = 0$ and minimal at $t = \frac{1}{2}$, while $J_-^i(t)$ is maximal at $t = \frac{1}{2}$ and minimal at $t = 0$.

Decomposition of jumps:

Since works like [25] and [26] a standard method for dealing with stochastic differential equations perturbed by Lévy noise with regularly varying tails is the decomposition of the Lévy process into a purely discontinuous big jump part and an independent small jump part.

Definition 3.14. *Assume $\rho \in (0,1)$. Split L into the big jump part $\eta = (\eta_t)_{t\geqslant 0}$ with characteristic triplet $\left(0, 0, \nu\left(\cdot \cap \left\{x \in \mathbb{R}^d : \|x\| \geqslant \varepsilon^{-\rho}\right\}\right)\right)$ and the small jump part $\xi = (\xi_t)_{t\geqslant 0} = (L_t - \eta_t)_{t\geqslant 0}$ that includes the continuous part of L and belongs to the triplet $\left(a, \Sigma, \nu\left(\cdot \cap \left\{x \in \mathbb{R}^d : \|x\| \in (0, \varepsilon^{-\rho})\right\}\right)\right)$.*

Definition 3.15. *Let $(\tau_n)_{n\in\mathbb{N}_0}$ with $\tau_0 = 0$ denote the sequence of jump times of η, $T_n = \tau_n - \tau_{n-1}$ is the inter-jump time and the jumps at $t = \tau_n$ are abbreviated by W_n. The scaled jump time $\frac{\tau_n}{2T}$ is abbreviated by $\hat{\tau}_n$.*

Remark 3.16. The random variable $\hat{\tau}_n - \hat{\tau}_{n-1}$ obeys an exponential distribution with parameter $2T\beta^\varepsilon$ where $\beta^\varepsilon = \nu\left(\left\{x \in \mathbb{R}^d : \|x\| \geqslant \varepsilon^{-\rho}\right\}\right) = \varepsilon^{\rho\alpha} l(\varepsilon^{-\rho})$ denotes the intensity of the compound Poisson process η and W_n admits the law $\frac{1}{\beta^\varepsilon}\nu\left(\cdot \cap \left\{x \in \mathbb{R}^d : \|x\| \geqslant \varepsilon^{-\rho}\right\}\right)$. The jump size is independent of the jump time.

3.3 The small jump process

The mathematical intuition already suggests that the big jump part of L is the main cause of exits of X^ε from a certain well. The aim of this section is to ensure that the contribution of the small jump process X^ξ to the essential behaviour of X^ε is really marginal because it rarely leaves the neighbourhood of the deterministic solution.

Definition 3.17. *For $s \geqslant 0$ define $\xi_t^s = \xi_{s+t} - \xi_s$, $t \geqslant 0$. The process $(X_{s,t}^\xi(x_0))_{t\geqslant 0}$ is called small jump process if it solves*

$$X_{s,t}^\xi(x_0) = x_0 - \int_0^t \nabla U\left(X_{s,r}^\xi, \frac{s+r}{2T}\right) \mathrm{d}r + \varepsilon \int_0^t g\left(X_{s,r-}^\xi\right) \mathrm{d}\xi_r^s, \quad t \geqslant 0. \tag{3.6}$$

Its existence and uniqueness follow as in the proof of Proposition 2.39. The deterministic solution of $\dot{x}(t) = -\nabla U(x(t), \frac{t}{2T})$ for $t \geqslant s$ with $x(s) = x_0$ was denoted by $x(t; x_0, s)$, $t \geqslant s$. Define $X^0_{s,t}(x_0) \coloneqq x(s+t; x_0, s)$ for $t \geqslant 0$ to embed it in the current notation.

Theorem 3.18. *Let $R \geqslant R^*$ and T_1 be exponentially distributed with parameter $\beta_\varepsilon = \nu(\|x\| \geqslant \varepsilon^{-\rho})$ and independent of ξ. Then there exist $\varepsilon_0, \gamma_0, \beta_0, p_0 > 0$ such that the estimate*

$$\sup_{s\geqslant 0} \sup_{x_0 \in \Omega^{\varepsilon^\gamma}_{\pm,R}(s)} \mathbb{P}\left(\sup_{t \leqslant T_1} \left\| X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0) \right\| \geqslant \varepsilon^\beta \right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

is valid for all $\varepsilon \in (0, \varepsilon_0), \gamma \in (0, \gamma_0), \beta \in (0, \beta_0)$ and $p \in (0, p_0)$.

The whole section is dedicated to the proof of this result following the methods of [38]. We spare the reader repeated referring to single parts of [38].
In a preparatory subsection we verify the boundedness of X^ξ in probability until a polynomial time $\varepsilon^{-\theta}$. Then we focus on the course of the trajectories until the logarithmic return time R_{ε^γ} defined in (Log) and later the behaviour of X^ξ near the well bottoms is analysed. This suffices to control the deviation of X^ξ and the deterministic process until a polynomial time $\varepsilon^{-\theta}$ which enables us to prove Theorem 3.18. Afterwards the arguments are transferred to the Marcus stochastic differential equation.

3.3.1 Preliminaries and the boundedness of the solution

Lemma 3.19. *There are $\varepsilon_0, \beta_0, \theta_0, p_0 > 0$ such that the jump part $[\xi]^d_t = \sum_{0<s\leqslant t} \|\Delta\xi_s\|^2$ of the quadratic variation process satisfies*

$$\mathbb{P}(\varepsilon^2 \, [\xi]^d_{\varepsilon^{-\theta}} \geqslant \varepsilon^\beta) \leqslant \mathrm{e}^{-\varepsilon^{-p}},$$

for $\varepsilon \in (0, \varepsilon_0), \beta \in (0, \beta_0), \theta \in (0, \theta_0)$ and $p \in (0, p_0)$.

Proof. From the Markov inequality we can derive

$$\mathbb{P}(\varepsilon^2 \, [\xi]^d_{\varepsilon^{-\theta}} \geqslant \varepsilon^\beta) = \mathbb{P}(\varepsilon^{2-2\beta} \, [\xi]^d_{\varepsilon^{-\theta}} \geqslant \varepsilon^{-\beta}) \leqslant \mathrm{e}^{-\varepsilon^{-\beta}} \mathbb{E}\, \mathrm{e}^{\varepsilon^{2-2\beta} [\xi]^d_{\varepsilon^{-\theta}}}.$$

It suffices to prove the boundedness of $\mathbb{E}\, \mathrm{e}^{\varepsilon^{2-2\beta}[\xi]^d_{\varepsilon^{-\theta}}}$ for $\varepsilon \to 0$. Since $[\xi]^d$ is an increasing Lévy process, a subordinator ([44], Proposition 3.11), its Laplace transform admits the following representation (Remark 21.6 in [42])

$$\mathbb{E}\, \mathrm{e}^{-\lambda [\xi]^d_t} = \exp\left(t \int_{\|x\| \in (0, \varepsilon^{-\rho})} \left(\mathrm{e}^{-\lambda \|x\|^2} - 1 \right) \mathrm{d}\nu(x) \right), \ \lambda \geqslant 0.$$

Because certain values of $\lambda < 0$ are also admissible, insert $\lambda = -\varepsilon^{2-2\beta}$ to get

$$\begin{aligned}
\mathbb{E}\, \mathrm{e}^{\varepsilon^{2-2\beta}[\xi]^d_{\varepsilon^{-\theta}}} &= \exp\left(\int\limits_{\|x\| \in (0,1)} \varepsilon^{-\theta} \big(\mathrm{e}^{\varepsilon^{2-2\beta}\|x\|^2} - 1 \big) \mathrm{d}\nu(x) + \int\limits_{\|x\| \in [1, \varepsilon^{-\rho})} \varepsilon^{-\theta} \big(\mathrm{e}^{\varepsilon^{2-2\beta}\|x\|^2} - 1 \big) \mathrm{d}\nu(x) \right) \\
&\leqslant \exp\left(2\, \varepsilon^{2-2\beta-\theta} \int\limits_{\|x\| \in (0,1)} \|x\|^2 \mathrm{d}\nu(x) + \varepsilon^{-\theta} \left(\mathrm{e}^{\varepsilon^{2-2\rho-2\beta}} - 1 \right) \nu\big(\{ \|x\| \geqslant 1 \} \big) \right),
\end{aligned}$$

because $\mathrm{e}^x - 1 \leqslant 2x$ for $x \in [0,1]$. For $0 < a < b$ the small noise limit of $\varepsilon^{-a}(\mathrm{e}^{\varepsilon^b} - 1)$ is zero. Thus for $2 - 2\rho - 2\beta - \theta > 0$ the exponent above is a null sequence as ε tends to zero which causes the boundedness of the expectation. □

Lemma 3.20. *Let $s \geqslant 0$ and M^s be given by $M_t^s = \xi_{s+t} - \xi_s - t\,\mathbb{E}\xi_1$. The process $(X_t)_{t\geqslant 0}$ denotes a d-dimensional, $(\mathcal{F}_{s+t})_{t\geqslant 0}$-adapted, a.s. bounded, and càdlàg process. There are $\varepsilon_0, p_0, \beta_0, \theta_0 > 0$ such that*

$$\sup_{s\geqslant 0} \mathbb{P}\left(\sup_{t\leqslant \varepsilon^{-\theta}} \varepsilon \Big| \sum_{j=1}^{d} \int_0^t X_{r-}^j \,\mathrm{d}M_r^{s,j} \Big| \geqslant \varepsilon^{\beta} \right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

holds for all $\varepsilon \in (0, \varepsilon_0), \beta \in (0, \beta_0), \theta \in (0, \theta_0)$ and $p \in (0, p_0)$.

Proof. The process M^s is a martingale (Theorems 32 and 41 in Chapter 1 of [39]). The property to be a local martingale is preserved under stochastic integration of adapted, càglàd processes (Theorem 29 in Chapter 3 of [39]) and thus

$$Y_t^s := \sum_{j=1}^{d} \int_0^t X_{r-}^j \,\mathrm{d}M_r^{s,j}$$

is a local martingale. The norm of its jumps is bounded by $c_1\,\varepsilon^{-\rho}$ for some $c_1 > 0$. Hence we can proceed as Kallenberg in the proof of Theorem 23.17 in [30]. Apply the following elementary estimate:

$$\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon|Y_t^s| \geqslant \varepsilon^{\beta} \right) \leqslant \mathbb{P}\left(\varepsilon^2 [Y^s]_{\varepsilon^{-\theta}} \geqslant \varepsilon^{4\beta} \right) + \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon|Y_t^s| \geqslant \varepsilon^{\beta}, \varepsilon^2[Y^s]_{\varepsilon^{-\theta}} < \varepsilon^{4\beta} \right).$$

Define $f(x) := -(x + \log(1-x)_+)x^{-2}$, $x \in \mathbb{R}$. Paraphrasing the method of Kallenberg yields

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon Y_t^s \geqslant \varepsilon^{\beta}, \varepsilon^2[Y^s]_{\varepsilon^{-\theta}} < \varepsilon^{4\beta} \Big) \\
&= \mathbb{P}\Big(\sup_{t\leqslant\varepsilon^{-\theta}} \mathrm{e}^{\varepsilon^{1-2\beta} Y_t^s} \geqslant \mathrm{e}^{\varepsilon^{-\beta}}, \varepsilon^2[Y^s]_{\varepsilon^{-\theta}} < \varepsilon^{4\beta} \Big) \\
&\leqslant \mathbb{P}\Big(\sup_{t\leqslant\varepsilon^{-\theta}} \exp\left(\varepsilon^{1-2\beta} Y_t^s + \varepsilon^{-4\beta} f(c_1\varepsilon^{1-2\beta-\rho})(\varepsilon^{4\beta} - \varepsilon^2[Y^s]_t) \right) \geqslant \mathrm{e}^{\varepsilon^{-\beta}}, \varepsilon^2[Y^s]_{\varepsilon^{-\theta}} < \varepsilon^{4\beta} \Big) \\
&\leqslant \mathbb{P}\Big(\sup_{t\leqslant\varepsilon^{-\theta}} \exp\left(\varepsilon^{1-2\beta} Y_t^s - \varepsilon^{2-4\beta} f(c_1\varepsilon^{1-2\beta-\rho})[Y^s]_t \right) \geqslant \mathrm{e}^{\varepsilon^{-\beta} - f(c_1\varepsilon^{1-2\beta-\rho})} \Big).
\end{aligned}$$

This transformation creates a supermartingale $\exp(\varepsilon^{1-2\beta} Y_t^s - \varepsilon^{2-4\beta} f(c_1\varepsilon^{1-2\beta-\rho})[Y^s]_t)$. On account of the useful inequality $C\mathbb{P}\left(\sup_{t\geqslant 0} Z_t \geqslant C\right) \leqslant \mathbb{E}Z_0$ for nonnegative supermartingales Z and $C \geqslant 0$ we obtain

$$\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon Y_t^s \geqslant \varepsilon^{\beta}, \varepsilon^2[Y^s]_{\varepsilon^{-\theta}} < \varepsilon^{4\beta} \right) \leqslant \mathrm{e}^{-\varepsilon^{-\beta} + f(c_1\varepsilon^{1-2\beta-\rho})} < \mathrm{e}^{-\varepsilon^{-q}},$$

because $f(c_1\varepsilon^{1-2\beta-\rho})$ converges as $\varepsilon \to 0$ if $1 - 2\beta - \rho > 0$ and therefore the upper bound $\mathrm{e}^{-\varepsilon^{-q}}$ is appropriate for some $q < \beta$. It is left to estimate the probability that the quadratic variation of εY^s until $t=\varepsilon^{-\theta}$ attains values bigger than $\varepsilon^{4\beta}$. The same methods as in step three of the proof of Proposition 2.39 verifies

$$\varepsilon^2 [Y^s]_{\varepsilon^{-\theta}} \leqslant c_2 \left(\varepsilon^{2-\theta} + \varepsilon^2 \sum_{t \leqslant \varepsilon^{-\theta}} \|\Delta \xi_t^s\|^2 \right),$$

for some $c_2 > 0$. Because the jump part of the quadratic variation of a Lévy process is a Lévy process itself (Proposition 3.11 in [44]), it is left to verify

$$\mathbb{P}(\varepsilon^2 \, [\xi]^d_{\varepsilon^{-\theta}} \geqslant \varepsilon^\gamma) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

for small $\varepsilon, \gamma, \theta$ and p. These arguments attribute the current lemma to the previous one. $\square$

The next lemma provides us with a boundedness result of the solution X^ξ.

Lemma 3.21. *For all $R \geqslant R^*$ there exist $N \in \mathbb{N}$ and $\varepsilon_0, \theta_0, p_0 > 0$ such that*

$$\sup_{s \geqslant 0} \sup_{x_0 \in O_R} \mathbb{P}\left(\sup_{t \leqslant \varepsilon^{-\theta}} \left\| X^\xi_{s,t}(x_0) \right\| \geqslant N \right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

is true for all $\varepsilon \in (0,\varepsilon_0), \theta \in (0,\theta_0)$ and $p \in (0,p_0)$.

Proof. Assume $s \geqslant 0$ and $x_0 \in O_R$. Remember the sequence $(u_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}_+$ appearing in the third step of the proof of Proposition 2.39. It fulfills $u_n \nearrow \infty$ as $n \to \infty$ and $\log(1+U(x,t)) \geqslant u_n$ for $x \in B_n^c(0)$, $t \geqslant 0$. Choose $N \in \mathbb{N}$ such that the term $\log\left(1+U\left(x_0, \frac{s}{2T}\right)\right)$ is smaller than $\frac{u_N}{2}$ for all $s \geqslant 0$ and $x_0 \in O_R$. The inclusion

$$\left\{ \sup_{t \leqslant \varepsilon^{-\theta}} \left\| X^\xi_{s,t}(x_0) \right\| \geqslant N \right\} \subseteq \left\{ \sup_{t \leqslant \varepsilon^{-\theta}} \log\left(1 + U\left(X^\xi_{s,t}(x_0), \frac{s+t}{2T} \right) \right) \geqslant u_N \right\}$$

is obvious. Applying the Itô formula for the function $f(x,t) = \log(1+U(x,t))$ yields

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t \leqslant \varepsilon^{-\theta}} \left\| X^\xi_{s,t}(x_0) \right\| \geqslant N \right) \\
&\leqslant \mathbb{P}\left(\log\left(1+U\left(x_0, \frac{s}{2T}\right)\right) \geqslant \frac{u_N}{2} \right) + \mathbb{P}\left(\sup_{t \leqslant \varepsilon^{-\theta}} \left| \int_0^t \frac{1}{2T} \frac{U_t(X^\xi_{s,r}, \frac{s+r}{2T})}{1+U(X^\xi_{s,r}, \frac{s+r}{2T})} \, \mathrm{d}r \right| \geqslant \varepsilon^\beta \right) \\
&\quad + \mathbb{P}\left(\sup_{t \leqslant \varepsilon^{-\theta}} \varepsilon \left| \int_0^t \frac{\nabla U(X^\xi_{s,r-}, \frac{s+r}{2T}) g(X^\xi_{s,r-})}{1+U(X^\xi_{s,r-}, \frac{s+r}{2T})} \, \mathrm{d}\xi_r^s \right| \geqslant \varepsilon^\beta \right) \qquad (3.7) \\
&\quad + \mathbb{P}\left(\sup_{t \leqslant \varepsilon^{-\theta}} \frac{\varepsilon^2}{2} \left| \sum_{i,j,k,l=1}^d \int_0^t f_{x_i x_j}\left(X^\xi_{s,r}, \frac{s+r}{2T} \right) g_{ik}(X^\xi_{s,r}) \, g_{jl}(X^\xi_{s,r}) \, \mathrm{d}[\xi^{s,k}, \xi^{s,l}]^c_r \right| \geqslant \varepsilon^\beta \right) \\
&\quad + \mathbb{P}\left(\sup_{t \leqslant \varepsilon^{-\theta}} \left| \sum_{r \leqslant t} \left[f\left(X^\xi_{s,r}, \frac{s+r}{2T} \right) - f\left(X^\xi_{s,r-}, \frac{s+r}{2T} \right) - \nabla f\left(X^\xi_{s,r-}, \frac{s+r}{2T} \right) \Delta X^\xi_{s,r} \right] \right| \geqslant \varepsilon^\beta \right).
\end{aligned}$$

Due to the well-considered choice of N the first summand is zero. The second probability also vanishes if $\alpha - \beta - \theta > 0$ and ε is small because $\frac{U_t}{1+U}$ is a bounded function according to (U4) and $(2T)^{-1}$ is of order ε^α since (T) is valid. An upper bound for the penultimate summand can be deduced from (U4) as in step three of the proof of Proposition 2.39

$$\mathbb{P}\bigg(\sup_{t\leqslant\varepsilon^{-\theta}}\frac{\varepsilon^2}{2}\Big|\sum_{i,j,k,l=1}^{d}\int_0^t f_{x_ix_j}\Big(X^\xi_{s,r},\frac{s+r}{2T}\Big)g_{ik}(X^\xi_{s,r})\,g_{jl}(X^\xi_{s,r})\,\mathrm{d}[\xi^{s,k},\xi^{s,l}]^c_r\Big|\geqslant\varepsilon^\beta\bigg)$$
$$\leqslant\mathbb{P}\Big(\sup_{t\leqslant\varepsilon^{-\theta}}c_1\,\varepsilon^2\,t\geqslant\varepsilon^\beta\Big)=0,$$

for some $c_1>0$ and $2-\theta-\beta>0$ and sufficiently small ε. The stated proof also helps to justify that the last summand of the right-hand side of (3.7) admits the upper bound $\mathbb{P}\left(c_2\varepsilon^2[\xi^s]^d_{\varepsilon^{-\theta}}\geqslant\varepsilon^\beta\right)$ for some $c_2>0$. This bound equals to the s-independent term $\mathbb{P}\left(c_2\varepsilon^2[\xi]^d_{\varepsilon^{-\theta}}\geqslant\varepsilon^\beta\right)$ due to the Lévy property of $[\xi]^d$. The exponential smallness of this expression emerges from the preparatory Lemma 3.19. It remains to handle the third summand which again can be splitted into an integral with respect to the Lebesgue measure and one with respect to the martingale $M^s_t:=\xi^s_t-t\,\mathbb{E}\xi_1$. Omitting the arguments of all functions we get

$$\mathbb{P}\bigg(\sup_{t\leqslant\varepsilon^{-\theta}}\varepsilon\Big|\int_0^t\frac{\nabla Ug}{1+U}\,\mathrm{d}\xi^s\Big|\geqslant\varepsilon^\beta\bigg)$$
$$\leqslant\mathbb{P}\bigg(\sup_{t\leqslant\varepsilon^{-\theta}}\varepsilon\Big|\int_0^t\frac{\nabla Ug\,\mathbb{E}\xi_1}{1+U}\,\mathrm{d}r\Big|\geqslant\varepsilon^{2\beta}\bigg)+\mathbb{P}\bigg(\sup_{t\leqslant\varepsilon^{-\theta}}\varepsilon\Big|\int_0^t\frac{\nabla Ug}{1+U}\,\mathrm{d}M^s\Big|\geqslant\varepsilon^{2\beta}\bigg).$$

The first probability converges to zero if $1-\rho-2\beta>0$ and Lemma 3.20 implies the exponential smallness of the second summand as ε converges to zero. □

3.3.2 Behaviour until the logarithmic return time and near the well minima

Lemma 3.22. *Assume $R\geqslant R^*$. There are $\varepsilon_0,\gamma_0,\beta_0,p_0>0$ such that the following inequality*

$$\sup_{s\geqslant0}\sup_{x_0\in\Omega^{\varepsilon^\gamma}_{\pm,R}(s)}\mathbb{P}\left(\sup_{t\leqslant R_{\varepsilon^\beta}}\left\|X^\xi_{s,t}(x_0)-X^0_{s,t}(x_0)\right\|\geqslant\varepsilon^\beta\right)\leqslant\mathrm{e}^{-\varepsilon^{-p}}$$

is valid for all $\varepsilon\in(0,\varepsilon_0),\gamma\in(0,\gamma_0),\beta\in(0,\beta_0)$ and $p\in(0,p_0)$.

Proof. Choose $N\in\mathbb{N}$ according to the previous lemma. Intersection with the event $B:=\left\{\sup_{t\leqslant R_{\varepsilon^\beta}}\|X^\xi_{s,t}(x_0)\|<N\right\}$ and its complement yields

$$\mathbb{P}\left(\sup_{t\leqslant R_{\varepsilon^\beta}}\left\|X^\xi_{s,t}(x_0)-X^0_{s,t}(x_0)\right\|\geqslant\varepsilon^\beta\right)$$
$$\leqslant\mathbb{P}\left(B^c\right)+\mathbb{P}\left(B\cap\left\{\sup_{t\leqslant R_{\varepsilon^\beta}}\left\|X^\xi_{s,t}(x_0)-X^0_{s,t}(x_0)\right\|\geqslant\varepsilon^\beta\right\}\right).$$

Since the logarithmic return time R_{ε^β} is much shorter than any polynomial time $\varepsilon^{-\theta}$ and a small noise parameter, the exponential smallness in ε of $\mathbb{P}(B^c)$ is known from the Lemma

3.21. The deviation of the X^ξ and the X^0 on the set B can be estimated as

$$\big\|X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\big\| \leqslant C_{U,N}\int_0^t \big\|X^\xi_{s,r}(x_0) - X^0_{s,r}(x_0)\big\|\mathrm{d}r + \varepsilon\Big\|\int_0^t g(X^\xi_{s,r-})\,\mathrm{d}\xi^s_r\Big\|$$

because ∇U is locally Lipschitz with Lipschitz constant $C_{U,N} > 0$ and both arguments are bounded on B. From Gronwall's inequality (Lemma 6.1 in [1]) we can derive

$$\big\|X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\big\| \leqslant \varepsilon\Big\|\int_0^t g(X^\xi_{s,r-})\,\mathrm{d}\xi^s_r\Big\| + \int_0^t \Big\|\int_0^r g(X^\xi_{s,u-})\,\mathrm{d}\xi^s_u\Big\| C_{U,N}\,\mathrm{e}^{C_{U,N}(t-r)}\mathrm{d}r,$$

for all $t \leqslant R_{\varepsilon^\beta}$. Taking the supremum of both sides yields

$$\sup_{t\leqslant R_{\varepsilon^\beta}} \big\|X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\big\| \leqslant \mathrm{e}^{C_{U,N}R_{\varepsilon^\beta}} \sup_{t\leqslant R_{\varepsilon^\beta}} \varepsilon\Big\|\int_0^t g(X^\xi_{s,r-})\,\mathrm{d}\xi^s_r\Big\|.$$

Since $\mathrm{e}^{C_{U,N}R_{\varepsilon^\beta}} = \mathrm{e}^{C_{U,N}\beta c|\log\varepsilon|} = \varepsilon^{-cC_{U,N}\beta}$, it remains to deal with

$$\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon\Big\|\int_0^t g(X^\xi_{s,r-})\,\mathrm{d}\xi^s_r\Big\| \geqslant \varepsilon^{\bar\beta}\right),$$

for $\bar\beta := \beta(1 + cC_{U,N})$. Again through passing to the martingale $M^s_t := \xi^s_t - t\,\mathbb{E}\xi_1$ we get

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon\Big\|\int_0^t g(X^\xi_{s,r-})\,\mathrm{d}\xi^s_r\Big\| \geqslant \varepsilon^{\bar\beta}\right)\\
&\leqslant \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon\Big\|\int_0^t g(X^\xi_{s,r-})\,\mathbb{E}\xi_1\,\mathrm{d}r\Big\| \geqslant \varepsilon^{2\bar\beta}\right) + \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon\Big\|\int_0^t g(X^\xi_{s,r-})\,\mathrm{d}M^s_r\Big\| \geqslant \varepsilon^{2\bar\beta}\right).
\end{aligned}$$

Apply the equivalence of all norms in $\mathbb{R}^d$, the estimate $\|\mathbb{E}\xi_1\| = \mathrm{O}(\varepsilon^{-\rho})$ (proof of Proposition 2.39), boundedness of all components of g and Lemma 3.20 to finish the proof. □

Lemma 3.23. *There exist constants $\varepsilon_0, \theta_0, \beta_0, p_0 > 0$ such that*

$$\sup_{s\geqslant 0}\ \sup_{x_0:\|x_0 - m_\pm(\frac{s}{2T})\|\leqslant\varepsilon^{2\beta}} \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \Big\|X^\xi_{s,t}(x_0) - m_\pm\Big(\frac{s+t}{2T}\Big)\Big\| \geqslant \varepsilon^\beta\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

holds for all $\varepsilon \in (0,\varepsilon_0), \theta \in (0,\theta_0), \beta \in (0,\beta_0)$ and $p \in (0, p_0)$.

Proof. The double-well shape of U was made precise through assumption (U2). The eigenvalues of $\left(\frac{\partial^2}{\partial x_i \partial x_j} U\left(m_\pm(t), t\right)\right)_{i,j=1}^d$ are positive and bounded away from zero. Therefore for x taken from a fixed neighbourhood of the well bottom $m_\pm(t)$ the relation

$$c\|x - m_\pm(t)\|^2 \leqslant U(x,t) - U(m_\pm(t), t) \leqslant C\|x - m_\pm(t)\|^2,$$

is true for some $C, c > 0$. If the small jump process leaves the ε^β-tube surrounding one minimum, it still lies within its vicinity at least for a short time due to the ε-dependent boundedness of the jumps of X^ξ. This justifies for x_0 with $\|x_0 - m_\pm(\frac{s}{2T})\| \leqslant \varepsilon^{2\beta}$

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \Big\|X^\xi_{s,t}(x_0) - m_\pm\Big(\frac{s+t}{2T}\Big)\Big\| \geqslant \varepsilon^\beta\right)\\
&\leqslant \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \left[U\Big(X^\xi_{s,t}(x_0), \frac{s+t}{2T}\Big) - U\Big(m_\pm\Big(\frac{s+t}{2T}\Big), \frac{s+t}{2T}\Big)\right] \geqslant c\,\varepsilon^{2\beta}\right).
\end{aligned}$$

As in the last lemma, intersection with $\left\{\sup_{t\leqslant R_{\varepsilon^\beta}} \|X^\xi_{s,t}(x_0)\| < N\right\}$ and its complement is advisable. The complement can be handled according to Lemma 3.21 and for a bounded small jump diffusion again exploit the Itô formula which gives

$$\begin{aligned}
&U\Big(X^\xi_{s,t}(x_0), \frac{s+t}{2T}\Big) - U\Big(m_\pm\Big(\frac{s+t}{2T}\Big), \frac{s+t}{2T}\Big) \\
&= U\Big(x_0, \frac{s}{2T}\Big) - U\Big(m_\pm\Big(\frac{s+t}{2T}\Big), \frac{s+t}{2T}\Big) \\
&\quad + \frac{1}{2T}\int_0^t U_t\Big(X^\xi_{s,r}, \frac{s+r}{2T}\Big)\mathrm{d}r - \int_0^t \Big\|\nabla U\Big(X^\xi_{s,r}, \frac{s+r}{2T}\Big)\Big\|^2 \mathrm{d}r \\
&\quad + \varepsilon\int_0^t \nabla U\Big(X^\xi_{s,r-}, \frac{s+r}{2T}\Big) g\Big(X^\xi_{s,r-}\Big)\mathrm{d}\xi^s_r \\
&\quad + \frac{\varepsilon^2}{2}\sum_{i,j,k,l=1}^d \int_0^t U_{x_ix_j}\Big(X^\xi_{s,r}, \frac{s+r}{2T}\Big) g_{ik}\Big(X^\xi_{s,r}\Big) g_{jl}\Big(X^\xi_{s,r}\Big)\mathrm{d}[\xi^{s,k}, \xi^{s,l}]^c_r \\
&\quad + \sum_{r\leqslant t}\Big[U\Big(X^\xi_{s,r}, \frac{s+r}{2T}\Big) - U\Big(X^\xi_{s,r-}, \frac{s+r}{2T}\Big) - \nabla U\Big(X^\xi_{s,r-}, \frac{s+r}{2T}\Big)\Delta X^\xi_{s,r}\Big].
\end{aligned} \tag{3.8}$$

Since $\|x_0 - m_\pm(\frac{s}{2T})\| \leqslant \varepsilon^{2\beta}$ the difference $U(x_0, \frac{s}{2T}) - U(m_\pm(\frac{s}{2T}), \frac{s}{2T})$ lies within $[0, C\varepsilon^{4\beta}]$. The slowly varying geometry of U caused by $\frac{1}{2T} \approx \varepsilon^\alpha$ and the boundedness of U_t on compact sets imply $U\Big(m_\pm(s), \frac{s}{2T}\Big) - U\Big(m_\pm\Big(\frac{s+t}{2T}\Big), \frac{s+t}{2T}\Big) \leqslant c_1\varepsilon^{\alpha-\theta}$ for some $c_1 > 0$ and all $t \leqslant \varepsilon^{-\theta}$, $s \geqslant 0$. Thus $\mathbb{P}(\sup_{t\leqslant\varepsilon^{-\theta}} |U(x_0, \frac{s}{2T}) - U(m_\pm(\frac{s+t}{2T}), \frac{s+t}{2T})| \geqslant \varepsilon^{3\beta})$ vanishes if $\alpha - \theta - 3\beta > 0$. As usual we exploit the negativity of the integral of $-\|\nabla U\|^2$ occuring in the third line of formula (3.8). For the treatment of the remaining summands the intersection with $\left\{\sup_{t\leqslant R_{\varepsilon^\beta}} \|X^\xi_{s,t}(x_0)\| < N\right\}$ is decisive to get boundedness of all occuring derivatives of U. Follow the lines of the proofs of 2.39 and 3.21 to complete this one. □

3.3.3 The exponential estimate

The last two lemmata already enabled us to prove Theorem 3.18 with $\varepsilon^{-\kappa}$ for $\kappa > 0$ instead of an exponentially distributed random variable which is later a decisive ingredient of the proof of this theorem.

Lemma 3.24. *If $R \geqslant R^*$ and $\kappa > 0$, then there are positive constants $\varepsilon_0, \beta_0, p_0, \gamma_0$ such that*

$$\sup_{s\geqslant 0}\ \sup_{x_0\in\Omega^{\varepsilon^\gamma}_{\pm,R}(s)} \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\kappa}} \left\|X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\right\| \geqslant \varepsilon^\beta\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

is true for all $\varepsilon \in (0, \varepsilon_0), \gamma \in (0, \gamma_0), \beta \in (0, \beta_0),$ and $p \in (0, p_0)$.

Proof. **Step 1:** First of all a modified version of the assertion is proven. We claim the existence of $\varepsilon_0, \beta_0, p_0, \theta_0, \gamma_0 > 0$ such that

$$\sup_{s\geqslant 0}\ \sup_{x_0\in\Omega^{\varepsilon^\gamma}_{\pm,R}(s)} \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \left\|X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\right\| \geqslant \varepsilon^\beta\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

holds for all $\varepsilon \in (0,\varepsilon_0), \gamma \in (0,\gamma_0), \beta \in (0,\beta_0), \theta \in (0,\theta_0)$ and $p \in (0,p_0)$. Decompose the interval $[0,\varepsilon^{-\theta}]$ into $[0,R_{\varepsilon^{5\beta}}]$ and $[R_{\varepsilon^{5\beta}},\varepsilon^{-\theta}]$ and estimate

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \left\|X^{\xi}_{s,t}(x_0)-X^0_{s,t}(x_0)\right\| \geqslant \varepsilon^{\beta}\right)\\
&\leqslant \mathbb{P}\left(\sup_{t\leqslant R_{\varepsilon^{5\beta}}} \left\|X^{\xi}_{s,t}(x_0)-X^0_{s,t}(x_0)\right\| \geqslant \varepsilon^{5\beta}\right)\\
&\quad+\mathbb{P}\left(\sup_{t\leqslant R_{\varepsilon^{5\beta}}} \left\|X^{\xi}_{s,t}(x_0)-X^0_{s,t}(x_0)\right\| < \varepsilon^{5\beta},\ \sup_{t\leqslant\varepsilon^{-\theta}} \left\|X^{\xi}_{s,t}(x_0)-X^0_{s,t}(x_0)\right\| \geqslant \varepsilon^{\beta}\right)\\
&=: I_1+I_2.
\end{aligned}$$

The exponential smallness of I_1 follows from Lemma 3.22. If $\gamma<5\beta$ and $x_0\in\Omega^{\varepsilon^{\gamma}}_{\pm}(s)$, then $X^0_{s,R_{\varepsilon^{5\beta}}}(x_0)$ at least lies within $B_{\varepsilon^{3\beta}}\left(p_{\pm}\left(s+R_{\varepsilon^{5\beta}}\right)\right)$ and thus I_2 admits the upper bound

$$\begin{aligned}
I_2 \leqslant \mathbb{P}\Bigg(&\left\|X^{\xi}_{s,R_{\varepsilon^{5\beta}}}(x_0)-X^0_{s,R_{\varepsilon^{5\beta}}}(x_0)\right\| < \varepsilon^{5\beta},\\
&\sup_{t\leqslant\varepsilon^{-\theta}} \left\|X^{\xi}_{s+R_{\varepsilon^{5\beta}},t}(X^{\xi}_{s,R_{5\beta}}(x_0)) - m_{\pm}\left(\frac{s+R_{\varepsilon^{5\beta}}+t}{2T}\right)\right\| \geqslant \varepsilon^{2\beta}\Bigg)\\
&+\sup_{r\geqslant 0}\ \sup_{\bar{x}_0:\|\bar{x}_0-p_{\pm}(r)\|\leqslant\varepsilon^{3\beta}} \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}}\left\|X^0_{r,t}(\bar{x}_0)-p_{\pm}(r+t)\right\|\geqslant\varepsilon^{2\beta}\right)\\
&+\mathbb{P}\left(\sup_{t\leqslant 2T}\left\|m_{\pm}\left(\frac{t}{2T}\right)-p_{\pm}(t)\right\|\geqslant\varepsilon^{2\beta}\right).
\end{aligned}$$

From Lemma 2.1 and assumption (T) we know that the deviation of the periodic solutions and the minima is of order $\varepsilon^{\alpha}<\varepsilon^{2\beta}$ if $2\beta<\alpha$. Hence the last summand of the upper bound of I_2 is zero. Due to the exponential fast approach of the deterministic solution towards the periodic solutions also the second summand vanishes for small ε. On $\left\{\left\|X^{\xi}_{s,R_{\varepsilon^{5\beta}}}(x_0)-X^0_{s,R_{\varepsilon^{5\beta}}}(x_0)\right\|<\varepsilon^{5\beta}\right\}$ the norm $\left\|X^{\xi}_{s,R_{\varepsilon^{5\beta}}}(x_0)-m_{\pm}\left(\frac{s+R_{\varepsilon^{5\beta}}}{2T}\right)\right\|$ is smaller than $\varepsilon^{4\beta}$ if $4\beta<\alpha$. Consequently the Markov property of $(X^{\xi}_{s,t})_{t\geqslant 0}$ justifies

$$I_2 \leqslant \sup_{u\geqslant 0}\ \sup_{x_0:\,\|x_0-m_{\pm}(u/2T)\|\leqslant\varepsilon^{4\beta}} \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}}\left\|X^{\xi}_{u,t}(x_0)-m_{\pm}\Big(\frac{u+t}{2T}\Big)\right\|\geqslant\varepsilon^{2\beta}\right)\leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

for sufficiently small β,θ,p and ε due to Lemma 3.23.

Step 2: Let $\varepsilon_0,\beta_0,p_0,\theta_0,\gamma_0>0$ exist according to the first step. Now choose $\theta<\theta_0$ and $k_{\varepsilon}\in\mathbb{N}$ with $\varepsilon^{-\kappa}\leqslant k_{\varepsilon}\varepsilon^{-\theta}$ and define

$$B_i := \left\{\sup_{t\in[i\varepsilon^{-\theta},(i+1)\varepsilon^{-\theta}]}\left\|X^{\xi}_{s,t}(x_0)-X^0_{s+i\varepsilon^{-\theta},t-i\varepsilon^{-\theta}}(X^{\xi}_{s,i\varepsilon^{-\theta}}(x_0))\right\|<\varepsilon^{2\beta}\right\}$$

for $i\in\{0,1,\dots,k_{\varepsilon}-1\}$. We claim the validity of the helpful inclusion

$$\bigcap_{i=0}^{n-1} B_i \subseteq C_n := \left\{\sup_{t\leqslant n\varepsilon^{-\theta}}\left\|X^{\xi}_{s,t}(x_0)-X^0_{s,t}(x_0)\right\|<\varepsilon^{\beta}\right\},\quad n\in\mathbb{N}.$$

If $n = 1$, the assertion is obvious. Assume the inclusion holds for some $k \in \mathbb{N}$. If $k = 1$, then $X^0_{s,\varepsilon^{-\theta}}(x_0) \in B_{\varepsilon^{2\beta}}(p_-(s+\varepsilon^{-\theta})) \cup B_{\varepsilon^{2\beta}}(p_+(s+\varepsilon^{-\theta}))$ for $\gamma < \beta$, because of (Log) and $\varepsilon^{-\theta} > R_{\varepsilon^\beta}$. Assume $k > 1$ then on the set $\cap_{i=0}^{k-2} B_i$ the deterministic solution $X^0_{s+(k-1)\varepsilon^{-\theta},\varepsilon^{-\theta}}(X^\xi_{s,(k-1)\varepsilon^{-\theta}}(x_0))$ lies also within the $\varepsilon^{2\beta}$-neighbourhood of one periodic solution, because the starting value $X^\xi_{s,(k-1)\varepsilon^{-\theta}}(x_0)$ is far away from the separatrix, since it belongs to $B_{\varepsilon^\beta}(X^0_{s,(k-1)\varepsilon^{-\theta}}(x_0))$. Thus on $\cap_{i=0}^{k-1} B_i$ the random variable $X^\xi_{s,k\varepsilon^{-\theta}}(x_0)$ at least belongs to the $2\varepsilon^{2\beta}$-neighbourhood of p_- or p_+. For $t \in [k\varepsilon^{-\theta},(k+1)\varepsilon^{-\theta}]$ we obtain on $\cap_{i=0}^{k} B_i$ the upper bound

$$\begin{aligned}\left\|X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\right\| \leqslant \varepsilon^{2\beta} + \left\|X^0_{s,t}(x_0) - p_\pm(s+t)\right\| \\ + \left\|p_\pm(s+t) - X^0_{s+k\varepsilon^{-\theta},t-k\varepsilon^{-\theta}}(X^\xi_{s,k\varepsilon^{-\theta}}(x_0))\right\|.\end{aligned}$$

Exploit assumption (Log) again and use the exponential attractivity of the periodic solutions to justify the upper bound ε^β of $\left\|X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\right\|$.
The importance of the proved inclusion becomes apparent through the succeeding estimate

$$\begin{aligned}\mathbb{P}\left(\sup_{t\leqslant k_\varepsilon\varepsilon^{-\theta}} \left\|X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\right\| \geqslant \varepsilon^\beta\right) &\leqslant \mathbb{P}\left(\bigcup_{i=0}^{k_\varepsilon-1} B_i^c\right) \\ &= \mathbb{P}(B_0^c) + \sum_{j=1}^{k_\varepsilon-1} \mathbb{P}(B_0 \cap \ldots \cap B_{j-1} \cap B_j^c).\end{aligned}$$

The probability $\mathbb{P}(B_0^c)$ is smaller than $\mathrm{e}^{-\varepsilon^{-p}}$ for $\varepsilon, p, \gamma, \theta$ and β according to step one. The treatment of the remaining summands requires $\cap_{i=0}^{j-1} B_j \subseteq C_j$ and the Markov property of X^ξ. From these arguments we can derive

$$\begin{aligned}\mathbb{P}(B_0 \cap \ldots \cap B_{j-1} \cap B_j^c) \leqslant \mathbb{P}\Bigg(&\|X^\xi_{s,j\varepsilon^{-\theta}}(x_0) - X^0_{s,j\varepsilon^{-\theta}}(x_0)\| < \varepsilon^\beta, \\ &\sup_{t\in[j\varepsilon^{-\theta},(j+1)\varepsilon^{-\theta}]} \left\|X^\xi_{s,t}(x_0) - X^0_{s+j\varepsilon^{-\theta},t-j\varepsilon^{-\theta}}(X^\xi_{s,j\varepsilon^{-\theta}}(x_0))\right\| \geqslant \varepsilon^{2\beta}\Bigg) \\ \leqslant \sup_{r\geqslant 0} \sup_{\bar{x}_0:\|x_0-p_\pm(r)\|\leqslant 2\varepsilon^\beta} &\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \left\|X^\xi_{r,t}(\bar{x}_0) - X^0_{r,t}(\bar{x}_0)\right\| \geqslant \varepsilon^{2\beta}\right).\end{aligned}$$

With the help of the result of the first step of the proof we finish the proof. □

The way is paved to verify Theorem 3.18.

Proof. To use the previous assertion replace T_1 by $\varepsilon^{-\kappa}$ for some $\kappa > \alpha\rho$ and estimate

$$\begin{aligned}&\mathbb{P}\left(\sup_{t\leqslant T_1} \left\|X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\right\| \geqslant \varepsilon^\beta\right) \\ &\leqslant \mathbb{P}\left(T_1 \geqslant \varepsilon^{-\kappa}\right) + \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\kappa}} \left\|X^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\right\| \geqslant \varepsilon^\beta\right).\end{aligned}$$

The exponential smallness in ε of the last summand of the upper bound is already known. Inserting the distribution of T_1 induces

$$\mathbb{P}\left(T_1 \geqslant \varepsilon^{-\kappa}\right) = \mathrm{e}^{-\varepsilon^{\alpha\rho-\kappa}\, l(\varepsilon^{-\rho})}.$$

Due to the choice of κ the result follows. □

3.3.4 The small jump process of the Marcus stochastic differential equation

It is natural to be tempted to apply the same procedure to the small jump process belonging to (3.2) and we will see that this attempt is fruitful.

Definition 3.25. *For $s \geqslant 0$ define $\xi^s_t = \xi_{s+t} - \xi_s$, $t \geqslant 0$. The process $(Z^\xi_{s,t}(x))_{t\geqslant 0}$ is called small jump process if it solves*

$$Z^\xi_{s,t}(x_0) = x_0 - \int_0^t \nabla U\left(Z^\xi_{s,r}, \frac{s+r}{2T}\right)\mathrm{d}r + \varepsilon \int_0^t g(Z^\xi_{s,r}) \diamond \mathrm{d}\xi^s_r, \quad t \geqslant 0.$$

Theorem 3.26. *Let the random variable T_1 be independent of ξ and exponentially distributed with parameter $\beta^\varepsilon = \nu(\|x\| \geqslant \varepsilon^{-\rho})$. Assume $R \geqslant R^*$. There are contants $\varepsilon_0, \gamma_0, \beta_0, p_0 > 0$ such that*

$$\sup_{s\geqslant 0} \sup_{x_0 \in \Omega^{\varepsilon^\gamma}_{\pm,R}(s)} \mathbb{P}\left(\sup_{t\leqslant T_1} \left\|Z^\xi_{s,t}(x_0) - X^0_{s,t}(x_0)\right\| \geqslant \varepsilon^\beta\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

is true for all $\varepsilon \in (0,\varepsilon_0), \gamma \in (0,\gamma_0), \beta \in (0,\beta_0)$ and $p \in (0,p_0)$.

The proof of Theorem 3.26 requires modified versions of Lemma 3.21, 3.22, 3.23 and 3.24. On the one hand the notation here is much more exhausting, on the other hand the powerful chain rule is available.

Proof. Define the enlarged process $(\bar{Z}^\xi_{s,t}(\bar{x}_0))_{t\geqslant 0} = (\bar{Z}^\xi_{s,t}(x_0), \frac{s+t}{2T})_{t\geqslant 0}$ with $\bar{x}_0 = (x_0, 0)$ that fulfills

$$\bar{Z}^\xi_{s,t}(x_0) = \bar{x}_0 + \int_0^t f\left(\bar{Z}^\xi_{s,r}\right) \diamond \mathrm{d}\bar{\xi}^s_r,$$

while $\bar{\xi}^s_t = (\xi^s_t, t)^T$ and

$$f(\bar{x}_1,\ldots,\bar{x}_{d+1}) = \begin{pmatrix} \varepsilon g(\bar{x}_1,\ldots,\bar{x}_d) & -\nabla U(\bar{x}_1,\ldots,\bar{x}_{d+1}) \\ 0 & \frac{1}{2T}\end{pmatrix}, \ (\bar{x}_1,\ldots,\bar{x}_{d+1}) \in \mathbb{R}^{d+1}.$$

Lemma 3.21 *for the Marcus case:*
For $\psi(\bar{x}) := \log(1+U(\bar{x}))$, $\bar{x} \in \mathbb{R}^{d+1}$ we have to verify

$$\mathbb{P}\left(\sup_{t\leqslant \varepsilon^{-\theta}} \psi\left(\bar{Z}^\xi_{s,t}(x)\right) \geqslant u_N\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

to prove the lemma. In the second step of the proof of Proposition 2.42 we already dealed with such an expression for the process $(\bar{Z}^{\xi,n}_{S,t})_{t\geqslant 0}$, while the index n originates from a truncation of the argument of the gradient ∇U and S denotes a stopping time. Transferring all arguments completes the current proof.

Lemma 3.22 *for the Marcus case:*
The main part of the proof is the verification of the estimate

$$\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}}\varepsilon\Big\|\int_0^t g(Z^\xi_{s,r})\diamond \mathrm{d}\xi^s_r\Big\|\geqslant\varepsilon^\beta\right)\leqslant \mathrm{e}^{-\varepsilon^{-p}}.$$

From the definition of the Marcus integral we derive

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}}\varepsilon\Big\|\int_0^t g(Z^\xi_{s,r})\diamond \mathrm{d}\xi^s_r\Big\|\geqslant\varepsilon^\beta\right)\\
&\leqslant \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}}\varepsilon\Big\|\int_0^t g(Z^\xi_{s,r-})\,\mathrm{d}\xi^s_r\Big\|\geqslant\varepsilon^{2\beta}\right)\\
&\quad+\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}}\frac{\varepsilon^2}{2}\Big\|\int_0^t g'(Z^\xi_{s,r})g(Z^\xi_{s,r})\,\mathrm{d}[\xi^s,\xi^s]^c_r\Big\|\geqslant\varepsilon^{2\beta}\right)\\
&\quad+\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}}\Big\|\sum_{r\leqslant t}\Big[\varphi(1;Z^\xi_{s,r-},\varepsilon\Delta\xi^s_r)-Z^\xi_{s,r-}-\varepsilon g(Z^\xi_{s,r-})\Delta\xi^s_r\Big]\Big\|\geqslant\varepsilon^{2\beta}\right)\\
&=: I_1+I_2+I_3.
\end{aligned}$$

Apply norm equivalence in $\mathbb{R}^d$, the boundedness of the components of g and Lemma 3.20 to treat I_1. From assumption (N2) we can deduce the upper bound $\varepsilon^2 c_1 t$, with $c_1>0$, for the norm of the integral with respect to the covariation process, which verifies $I_2=0$ for $2-2\beta-\theta>0$ and small ε. Remark 2.32 transfers the problem to Lemma 3.19.

Lemma 3.23 *for the Marcus case:*
As its Itô counterpart X^ξ the process Z^ξ only displays very small jumps. Thus the exponential smallness in ε of the term

$$\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}}\Big[U\Big(\bar Z^\xi_{s,t}(x_0)\Big)-U\Big(m_\pm\Big(\frac{s+t}{2T}\Big),\frac{s+t}{2T}\Big)\Big]\geqslant\varepsilon^\beta\right)$$

must be proven. Again use the chain rule for $U\big(\bar Z^\xi_{s,t}(x_0)\big)$ and follow the lines of the modified proof of Lemma 3.21. Be aware that here an intersection with the set $\{\sup_{t\leqslant\varepsilon^{-\theta}}\|\bar Z^\xi_{s,t}\|\leqslant N\}$ is decisive to get bounded integrands.

The Markov property of the solution of the Marcus stochastic differential equation, Lemmata 3.22 and 3.23 directly verify Lemma 3.24 which is the last ingredient of the proof of Theorem 3.26. □

3.4 Transition within the time interval $[T(1-\delta), T(1+\delta)]$

Roughly speaking a transition of X^ε_{2Tt} from one region of attraction to the other one is caused by a big jump W_1 with $m_\pm(t) + \varepsilon g(m_\pm(t))W_1 \in \Omega_\mp(2Tt)$. Motivated through the section about the approximation by two-state Markov chains we are going to analyse the probability that X^ε_t respectively Z^ε_t jump from one bounded and reduced domain to the other one in the time interval $[T(1-\delta), T(1+\delta)]$. The necessity of the consideration of bounded and reduced domains is caused by the uniform estimate of the probability of a large deviation of the small jump process from the deterministic solution for starting values taken from such domains (Theorem 3.18 respectively Theorem 3.26).

3.4.1 Definitions and important estimates

Frequently used definitions and estimates are added as prefix.

Definition 3.27. *(i) For $x_0 \in \mathbb{R}^d$ the process $(\hat{X}^\varepsilon_t(x_0))_{t\geqslant 0}$ is given by $\hat{X}^\varepsilon_t(x_0) := X^\varepsilon_{2Tt}(x_0)$ and for $s \geqslant 0$ define $\hat{X}^\xi_{s,t}(x_0) := X^\xi_{2Ts,2Tt}(x_0)$, $t \geqslant 0$. Remember the definition $\hat{\tau}_j = \frac{\tau_j}{2T}$ of the scaled jump times of the big jump part η of L.*

(ii) Let $R \geqslant R^*$ and $x_0 \in \Omega^{\varepsilon^\gamma}_{\pm,R}(0)$. Define the first exit time of $\hat{X}^\varepsilon_t$ from $\Omega^{\varepsilon^\gamma}_{\pm,R}(2Tt)$ by

$$\hat{\tau}^\varepsilon_\pm := \inf\left\{t \geqslant 0 : \hat{X}^\varepsilon_t(x_0) \notin \Omega^{\varepsilon^\gamma}_{\pm,R}(2Tt)\right\}.$$

Definition 3.28. *Assume $R \geqslant R^*$, $j \in \mathbb{N}_0$, $0 \leqslant s \leqslant t$ and $x \in \mathbb{R}$. Define the δ-inner part of a set B by $B^{-\delta} := \{x \in B : \operatorname{dist}(\partial B, x) \geqslant \delta\}$ for $\delta > 0$ and relevant events are given by*

$$\begin{aligned}
B^{\pm,j}_{s,t}(x) &= \Big\{\hat{X}^\xi_{s,r}(x) \in \Omega^{\varepsilon^\gamma}_{\pm,R}(2T(s+r)),\ r \in [0, t-s), \\
&\qquad \hat{X}^\xi_{s,t-s}(x) + \varepsilon g(\hat{X}^\xi_{s,t-s}(x))W_{j+1} \in \Omega^{\varepsilon^\gamma}_{\pm,R}(2Tt)\Big\}, \\
\bar{B}^{\pm,j}_{s,t}(x) &= \Big\{\hat{X}^\xi_{s,r}(x) \in \Omega^{\varepsilon^\gamma}_{\pm,R}(2T(s+r)),\ r \in [0, t-s), \\
&\qquad \hat{X}^\xi_{s,t-s}(x) + \varepsilon g(\hat{X}^\xi_{s,t-s}(x))W_{j+1} \in \Omega^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm}(2Tt) \cap O^{-\varepsilon^{2\gamma}}_R\Big\}, \\
C^{\pm,j}_{s,t}(x) &= \Big\{\hat{X}^\xi_{s,r}(x) \in \Omega^{\varepsilon^\gamma}_{\pm,R}(2T(s+r)),\ r \in [0, t-s), \\
&\qquad \hat{X}^\xi_{s,t-s}(x) + \varepsilon g(\hat{X}^\xi_{s,t-s}(x))W_{j+1} \in \Omega_{\mp,R}(2Tt)\Big\}, \\
E_{s,t}(x) &= \left\{\sup_{r\in[0,t-s]} \|\hat{X}^\xi_{s,r}(x) - X^0_{s,r}(x)\| \leqslant \frac{1}{3}\varepsilon^{4\gamma}\right\}, \\
F^\pm_{s,t}(x) &= \Big\{\exists u \in (0, t-s) : \hat{X}^\xi_{s,r}(x) \in \Omega^{\varepsilon^\gamma}_{\pm,R}(2T(s+r)),\ r \in [0, u), \\
&\qquad \hat{X}^\xi_{s,u}(x) \notin \Omega^{\varepsilon^\gamma}_{\pm,R}(2T(s+u))\Big\}.
\end{aligned}$$

In addition decisive space dependent jump sets are defined.

Definition 3.29. *For $R \geqslant R^*$, $\varepsilon, \gamma, t \geqslant 0$ and $x \in \mathbb{R}^d$ define*

$$
\begin{aligned}
D_{\mp,R}(x,t) &= \left\{ y \in \mathbb{R}^d : x + g(x)y \in \Omega_{\mp,R}(2Tt) \right\}, \\
D^{\varepsilon^\gamma}_{\mp,R}(x,t) &= \left\{ y \in \mathbb{R}^d : x + g(x)y \in \Omega^{\varepsilon^\gamma}_{\mp,R}(2Tt) \right\}, \\
D^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\mp,R}(x,t) &= \left\{ y \in \mathbb{R}^d : x + g(x)y \in \Omega^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\mp}(2Tt) \cap O_R^{-\varepsilon^{2\gamma}} \right\}.
\end{aligned}
$$

The following result is a helpful statement which allows to get rid of difficult stochastic expressions and instead deal with Lebesgue integrals and probabilities as integrands.

Lemma 3.30. *Assume $k \in \mathbb{N}\backslash\{1\}$, $I \subseteq \mathbb{R}_+$ is an interval and $0 \leqslant s \leqslant t$. Let $B_j(r)$, $r \geqslant 0$ for $j = 0, \dots, k$ be equal to $\Omega^{\varepsilon^\gamma}_{\pm,R}(2Tr)$, $\Omega^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm}(2Tr) \cap O_R^{-\varepsilon^{2\gamma}}$ or $\Omega_{\mp,R}(2Tr)$. For all $j = 0, \dots, k-1$ and $x \in \mathbb{R}^d$, let the event $S^j_{s,t}(x)$ admit the structure (cf. Definition 3.28)*

$$
S^j_{s,t}(x) = \left\{ \hat{X}^\xi_{s,r}(x) \in B_j(s+r),\, r \in [0, t-s),\, \hat{X}^\xi_{s,t-s}(x) + \varepsilon g(X^\xi_{s,t-s}(x)) W_{j+1} \in B_{j+1}(t) \right\}.
$$

Define $s_0 := 0$ and choose $x_0 \in B_0(s_0)$. Then the following estimate holds:

$$
\begin{aligned}
&\mathbb{E}\, \mathbf{1}_{\{\hat{\tau}_k \in I\}} \prod_{j=0}^{k-1} \mathbf{1}\left(S^j_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x_0)) \right) \\
&\geqslant \int_I \int_0^{s_k} \dots \int_0^{s_2} (2T\beta^\varepsilon)^k \, \mathrm{e}^{-2T\beta^\varepsilon s_k} \prod_{j=0}^{k-1} \inf_{x \in B_j(s_j)} \mathbb{P}(S^j_{s_j,s_{j+1}}(x)) \, \mathrm{d}s_1 \dots \mathrm{d}s_k.
\end{aligned}
$$

A similar upper bound is valid, where the infima are replaced by suprema.

Proof. The main tool of the proof is the strong Markov property of $(\hat{X}^\varepsilon_t, t)_{t\geqslant 0}$ (Proposition 2.43). Introduce the function

$$
h_{k-1}\left(\hat{X}^\varepsilon_{\hat{\tau}_{k-1}}(x_0), \hat{\tau}_{k-1} \right) := \mathbb{E}\left(\mathbf{1}_I(\hat{\tau}_k) \mathbf{1}\left(S^{k-1}_{\hat{\tau}_{k-1},\hat{\tau}_k}(\hat{X}^\varepsilon_{\hat{\tau}_{k-1}}(x_0)) \right) \Big| \mathcal{F}_{\tau_{k-1}} \right).
$$

Conditioning on $\mathcal{F}_{\tau_{k-1}}$ and the definition of the event $S^{k-2}_{\hat{\tau}_{k-2},\hat{\tau}_{k-1}}(x)$ yield

$$
\mathbb{E}\, \mathbf{1}_I(\hat{\tau}_k) \prod_{j=0}^{k-1} \mathbf{1}\left(S^j_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x_0)) \right) \geqslant \mathbb{E}\left(\prod_{j=0}^{k-2} \mathbf{1}\left(S^j_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x_0)) \right) \inf_{x \in B_{k-1}(\hat{\tau}_{k-1})} h_{k-1}(x, \hat{\tau}_{k-1}) \right).
$$

Iterate the application of the strong Markov property and successively define the functions

$$
h_j\left(\hat{X}^\varepsilon_{\hat{\tau}_j}(x_0), \hat{\tau}_j \right) := \mathbb{E}\left(\inf_{x \in B_{j+1}(\hat{\tau}_{j+1})} h_{j+1}(x, \hat{\tau}_{j+1})\, \mathbf{1}\left(S^j_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x_0)) \right) \Big| \mathcal{F}_{\tau_j} \right),
$$

for decreasing indices $j = k-2, \dots, 1$ until we have

$$
\mathbb{E}\, \mathbf{1}_I(\hat{\tau}_k) \prod_{j=0}^{k-1} \mathbf{1}\left(S^j_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x_0)) \right) \geqslant \mathbb{E}\left(\mathbf{1}\left(S^0_{0,\hat{\tau}_1}(x_0) \right) \inf_{x \in B_1(\hat{\tau}_1)} h_1(x, \hat{\tau}_1) \right).
$$

We use the independence of the jump times and sizes of the big jump part and the small jump process and the explicit density of $\hat{\tau}_1$ (Remark 3.16) to get

$$\begin{aligned}
&\mathbb{E}\left(\mathbf{1}\left(S^0_{0,\hat{\tau}_1}(x_0)\right)\inf_{x\in B_1(\hat{\tau}_1)} h_1(x,\hat{\tau}_1)\right)\\
&\geqslant \int_0^\infty 2T\beta^\varepsilon\, \mathrm{e}^{-2T\beta^\varepsilon t_1}\inf_{x_0\in B_0(0)}\mathbb{P}(S^0_{0,t_1}(x_0))\inf_{x_1\in B_1(t_1)} h_1(x_1,t_1)\,\mathrm{d}t_1.
\end{aligned}$$

The conditional law of the inter jump time $\hat{\tau}_{j+1}-\hat{\tau}_j$ given $\hat{\tau}_j = t_1+\ldots+t_j$ is an exponential distribution with parameter $2T\beta^\varepsilon$. The successive insertion of the exponential densities follows. Defining $t_0 = 0$ and interchanging integrals with infima results in

$$\begin{aligned}
\mathbb{E}\,\mathbf{1}_I(\hat{\tau}_k)\prod_{j=0}^{k-1}\mathbf{1}\left(S^j_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x_0))\right) &\geqslant \int_0^\infty\ldots\int_0^\infty (2T\beta^\varepsilon)^k\,\mathrm{e}^{-2T\beta^\varepsilon\sum_{j=1}^k t_j}\mathbf{1}\Big\{\sum_{j=1}^k t_j\in I\Big\}\\
&\prod_{j=0}^{k-1}\inf_{x\in B_j(\sum_{i=0}^j t_i)}\mathbb{P}(S^j_{\sum_{i=0}^j t_i,\sum_{i=0}^{j+1} t_i}(x))\,\mathrm{d}t_1\ldots\mathrm{d}t_k.
\end{aligned}$$

Substitute $(s_1,\ldots,s_k) = (t_1, t_1+t_2,\ldots,t_1+\ldots+t_k)$ to finish the proof of the lower bound. □

Inter-jump times which are bigger than the logarithmic return time and a small jump process that hardly deviates from the deterministic trajectory guarantee that the big jump of $\hat{X}^\varepsilon$ is performed from the vicinity of one well bottom. Rigorous estimates are given here.

Lemma 3.31. *Assume $c>0$, $j\in\mathbb{N}_0$, $R>R^*$. Then there are $\varepsilon_0,\gamma_0,p_0>0$ such that for all $0\leqslant t-s\leqslant c$ the following estimates hold true for all $\varepsilon\in(0,\varepsilon_0), p\in(0,p_0)$ and $\gamma\in(0,\gamma_0)$:*

(i)

$$\begin{aligned}
&\mathbf{1}_{\{2T(t-s)\geqslant R_{\varepsilon^\gamma}\}}\sup_{x\in\Omega^{\varepsilon^\gamma}_{\pm,R}(2Ts)}\mathbb{P}(B^{\pm,j}_{s,t}(x))\\
&\leqslant \mathrm{e}^{-\varepsilon^{-p}} + \sup_{x\colon\|x-m_\pm(t)\|\leqslant\varepsilon^{4\gamma}}\frac{1}{\beta^\varepsilon}\nu\left(B^c_{\varepsilon^{-\rho}}(0)\cap\frac{1}{\varepsilon}D^{\varepsilon^\gamma}_{\pm,R}(x,t)\right),
\end{aligned}$$

(ii)

$$\mathbf{1}_{\{2T(t-s)\geqslant R_{\varepsilon^\gamma}\}}\sup_{x\in\Omega^{\varepsilon^\gamma}_{\pm,R}(2Ts)}\mathbb{P}(C^{\pm,j}_{s,t}(x))\leqslant \mathrm{e}^{-\varepsilon^{-p}} + \sup_{x\colon\|x-m_\pm(t)\|\leqslant\varepsilon^{4\gamma}}\frac{1}{\beta^\varepsilon}\nu\left(\frac{1}{\varepsilon}D_{\mp,R}(x,t)\right),$$

(iii)

$$\begin{aligned}
&\mathbf{1}_{\{2T(t-s)\geqslant R_{\varepsilon^\gamma}\}}\inf_{x\in\Omega^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm}(2Ts)\cap O_R^{-\varepsilon^{2\gamma}}}\mathbb{P}(\bar{B}^{\pm,j}_{s,t}(x))\\
&\geqslant \left(1-\mathrm{e}^{-\varepsilon^{-p}}\right)\mathbf{1}_{\{2T(t-s)\geqslant R_{\varepsilon^\gamma}\}}\inf_{x\colon\|x-m_\pm(t)\|\leqslant\varepsilon^{4\gamma}}\frac{1}{\beta^\varepsilon}\nu\left(B^c_{\varepsilon^{-\rho}}(0)\cap\frac{1}{\varepsilon}D^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm,R}(x,t)\right),
\end{aligned}$$

(iv)

$$\begin{aligned}
&\mathbf{1}_{\{2T(t-s)\geqslant R_{\varepsilon^\gamma}\}}\inf_{x\in\Omega^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm}(2Ts)\cap O_R^{-\varepsilon^{2\gamma}}}\mathbb{P}(C^{\pm,j}_{s,t}(x))\\
&\geqslant \left(1-\mathrm{e}^{-\varepsilon^{-p}}\right)\mathbf{1}_{\{2T(t-s)\geqslant R_{\varepsilon^\gamma}\}}\inf_{x\colon\|x-m_\pm(t)\|\leqslant\varepsilon^{4\gamma}}\frac{1}{\beta^\varepsilon}\nu\left(\frac{1}{\varepsilon}D_{\mp,R}(x,t)\right).
\end{aligned}$$

Proof. Intersection with the event $E_{s,t}(x)$ that stands for a small deviation of the small jump process and the deterministic solution and its complement yields

$$\mathbb{P}(B^{\pm,j}_{s,t}(x)) \leqslant \mathbb{P}(E_{s,t}(x) \cap B^{\pm,j}_{s,t}(x)) + \mathbb{P}(E^c_{s,t}(x)).$$

Since $t-s \leqslant c$ and assumption (T) holds the inter-jump period $2T(t-s)$ is at most of order $\varepsilon^{-\alpha}$. Apply Lemma 3.24 to justify the uniform exponential smallness of $\mathbb{P}(E^c_{s,t}(x))$ in ε. If $2T(t-s) \geqslant R_{\varepsilon^\gamma}$ holds, $\alpha > 4\gamma$, and the small jump process lies in the $\frac{1}{3}\varepsilon^{4\gamma}$-neighbourhood of the deterministic solution, the random variable $\hat{X}^{\xi}_{s,t-s}(x)$ belongs to $B_{\varepsilon^{4\gamma}}(m_\pm(t))$. The Markov property and the knowledge of the law of the jump W_1 (Remark 3.16) allow us to write

$$\begin{aligned}
&\mathbf{1}_{\{2T(t-s)\geqslant R_{\varepsilon^\gamma}\}} \sup_{x\in\Omega^{\varepsilon^\gamma}_{\pm,R}(2Ts)} \mathbb{P}(E_{s,t}(x) \cap B^{\pm,j}_{s,t}(x)) \\
&\leqslant \sup_{x\colon \|x-m_\pm(t)\|\leqslant \varepsilon^{4\gamma}} \frac{1}{\beta^\varepsilon}\nu\left(B^c_{\varepsilon^{-\rho}}(0) \cap \frac{1}{\varepsilon} D^{\varepsilon^\gamma}_{\pm,R}(x,t)\right),
\end{aligned}$$

which verifies (i). The proof of (ii) demands similar arguments together with $\frac{1}{\varepsilon}D_{\mp,R}(x,t) \subseteq B^c_{\varepsilon^{-\rho}}(0)$. To justify (iii) again intersect with $E_{s,t}(x)$ which immediately guarantees that the small jump process never leaves the bounded and reduced basin $\Omega^{\varepsilon^\gamma}_{\pm,R}(2Tt)$. This implies

$$\mathbb{P}(\bar{B}^{\pm,j}_{s,t}(x)) \geqslant \mathbb{P}\Big(E_{s,t}(x) \cap \Big\{\hat{X}^{\xi}_{s,t-s}(x) + \varepsilon g(\hat{X}^{\xi}_{s,t-s}(x))W_{j+1} \in \Omega^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm}(2Tt) \cap O^{-\varepsilon^{2\gamma}}_{R}\Big\}\Big).$$

Make use of the inequality $2T(t-s) \geqslant R_{\varepsilon^\gamma}$, the independence of the small jump process ξ and η, and the Markov property to get

$$\begin{aligned}
&\mathbf{1}_{\{2T(t-s)\geqslant R_{\varepsilon^\gamma}\}} \inf_{x\in\Omega^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm}(2Ts)\cap O^{-\varepsilon^{2\gamma}}_{R}} \mathbb{P}(\bar{B}^{\pm,j}_{s,t}(x)) \\
&\geqslant \inf_{x\colon \|x-m_\pm(t)\|\leqslant \varepsilon^{4\gamma}} \frac{1}{\beta^\varepsilon}\nu\left(B^c_{\varepsilon^{-\rho}}(0) \cap \frac{1}{\varepsilon} D^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm,R}(x,t)\right)\mathbb{P}(E_{s,t}(x)).
\end{aligned}$$

The inequality $\mathbb{P}(E_{s,t}(x)) \geqslant 1 - \mathrm{e}^{-\varepsilon^{-p}}$ completes the proof of (iii). The proof of (iv) is analogous. ☐

Through a combination of Lemma 3.30, for example with $S^j_{s,t}(x) = B^{\pm,j}_{s,t}(x)$, and the previous lemma we should be prepared to handle iterated integrals of a special structure. Hence we formulate two simple geometric and one probabilistic result.

Lemma 3.32. *(i) Let $C, c > 0$ and $n \in \mathbb{N}$ be such that $C > nc > 0$. Then the Lebesgue volume of the set*

$$\{t \in [0,C]^n : t_1 \geqslant c, t_2 \geqslant t_1 + c, \ldots, t_n \geqslant t_{n-1} + c\}$$

equals to $\frac{(C-nc)^n}{n!}$.

(ii) Assume $C > 0, n \in \mathbb{N}$ and $f\colon [0,C] \to \mathbb{R}_+$ denotes an integrable function. Then it holds

$$\int_{\mathbb{R}_+} \cdots \int_{\mathbb{R}_+} \mathbf{1}_{\{0\leqslant t_1\leqslant\ldots\leqslant t_n\leqslant C\}} \prod_{j=1}^{n} f(t_j)\,\mathrm{d}t_1 \ldots \mathrm{d}t_n = \frac{1}{n!}\left(\int_0^C f(s)\,\mathrm{d}s\right)^n.$$

Proof. The calculation of the volume in (i) demands the evaluation of the iterated integral

$$\int_{nc}^{C}\int_{(n-1)c}^{t_n-c}\dots\int_{c}^{t_2-c} 1\,\mathrm{d}t_1\dots\mathrm{d}t_n.$$

By mathematical induction over n, the inner $n-1$ integrals represent the volume of the set

$$\left\{(t_1,\dots,t_{n-1})\in[0,t_n-c]^{n-1}: t_1\geqslant c, t_2\geqslant t_1+c,\dots,t_{n-1}\geqslant t_{n-2}+c\right\},$$

which is $\frac{(t_n-c-(n-1)c)^{n-1}}{(n-1)!}$. Integrating this expression over $[nc,C]$ yields assertion (i). The main ingredients of the proof of (ii) are the decomposition of $[0,C]^n$ into the sets $\{t\in[0,C]^n : t_{\pi(1)}\leqslant\dots\leqslant t_{\pi(n)}\}$ of equal size where π denotes one of the $n!$ permutations of $\{1,\dots,n\}$ and the invariance of the product $\prod_{j=1}^n f(t_{\pi(j)})$ with respect to a permutation π. □

Lemma 3.33. *Assume X is a Poisson distributed random variable with parameter $\lambda>0$. Then for all $k\in\mathbb{N}_0$ we have*

$$\mathbb{P}(X\geqslant k)\leqslant\frac{\lambda^k}{k!}.$$

Proof. The probability that X attains the value $n\in\mathbb{N}_0$ equals to $\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}$. Thus we obtain

$$\mathbb{P}(X\geqslant k)=\sum_{n=k}^{\infty}\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}=\mathrm{e}^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^{n+k}}{(n+k)!}=\frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\,\frac{n!k!}{(n+k)!}\leqslant\frac{\lambda^k}{k!}$$

while the inequality follows from $\frac{(n+k)!}{n!k!}=\binom{n+k}{n}$ and $\binom{n+k}{n}\geqslant 1$. □

Finally the space dependence of the jump sets $D_{\pm,R}(x,t)$, $D^{\varepsilon^\gamma}_{\pm,R}(x,t)$ and $D^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm,R}(x,t)$ can be omitted if several error terms are accepted as presented below.

Definition 3.34. *Assume $\delta>0$ and $B\in\mathcal{B}(\mathbb{R}^d)$ then $B^{+\delta}:=B\cup\left\{x\in\mathbb{R}^d:\operatorname{dist}(B,x)\leqslant\delta\right\}$ denotes the δ-enlargement of B. For $R\geqslant R^*$, $\varepsilon,\gamma,t\geqslant 0$ and $x\in\mathbb{R}^d$ introduce the sets*

$$\begin{aligned}
D_{\mp,R}(t)&=\Big\{y\in\mathbb{R}^d: m_\pm(t)+g(m_\pm(t))y\in\Omega_{\mp,R}(2Tt)\Big\},\\
G^{\varepsilon}_{\mp,R}(t)&=\Big\{y\in\mathbb{R}^d: m_\pm(t)+g(m_\pm(t))y\in\Gamma_R^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(2Tt)\cup\Big(O_R^{+\varepsilon^{2\gamma}}\backslash O_R^{-2\varepsilon^{2\gamma}}\Big)\Big\},\\
H^{\varepsilon}_{\mp,R}(t)&=\Big\{y\in\mathbb{R}^d: m_\pm(t)+g(m_\pm(t))y\in\Big(O_R^{-2\varepsilon^{2\gamma}}\Big)^c\Big\},\\
H_{\mp,R}(t)&=\Big\{y\in\mathbb{R}^d: m_\pm(t)+g(m_\pm(t))y\in O_R^c\Big\}.
\end{aligned}$$

Lemma 3.35. *Assume $R>R^*$, $\gamma>0$ and $t\geqslant 0$. The succeeding estimates hold*

(i) $$\sup_{x:\,\|x-m_\pm(t)\|\leqslant\varepsilon^{4\gamma}}\nu\Big(\frac{1}{\varepsilon}D_{\mp,R}(x,t)\Big)\leqslant\nu\Big(\frac{1}{\varepsilon}D_{\mp,R}(t)\Big)+\nu\Big(\frac{1}{\varepsilon}G^{\varepsilon}_{\mp,R}(t)\Big)+\nu(B^c_{\varepsilon^{-1-\gamma}}(0)),$$

(ii) $$\inf_{x:\,\|x-m_\pm(t)\|\leqslant\varepsilon^{4\gamma}}\nu\Big(\Big(\frac{1}{\varepsilon}D^{\varepsilon^\gamma}_{\pm,R}(x,t)\Big)^c\Big)\geqslant\nu\Big(\frac{1}{\varepsilon}D_{\mp,R}(t)\Big)-\nu(B^c_{\varepsilon^{-1-\gamma}}(0)),$$

(iii) $$\inf_{x:\,\|x-m_\pm(t)\|\leqslant\varepsilon^{4\gamma}}\nu\Big(\frac{1}{\varepsilon}D_{\mp,R}(x,t)\Big)\geqslant\nu\Big(\frac{1}{\varepsilon}D_{\mp,R}(t)\Big)-\nu\Big(\frac{1}{\varepsilon}G^{\varepsilon}_{\mp,R}(t)\Big)-\nu(B^c_{\varepsilon^{-1-\gamma}}(0)),$$

(iv) $$\begin{aligned}\sup_{x:\,\|x-m_\pm(t)\|\leqslant\varepsilon^{4\gamma}}\nu\Big(\Big(\frac{1}{\varepsilon}D^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm,R}(x,t)\Big)^c\Big)&\leqslant\nu\Big(\frac{1}{\varepsilon}D_{\mp,R}(t)\Big)+\nu\Big(\frac{1}{\varepsilon}G^{\varepsilon}_{\mp,R}(t)\Big)+\nu(B^c_{\varepsilon^{-1-\gamma}}(0))\\&\quad+\nu\Big(\frac{1}{\varepsilon}H^{\varepsilon}_{\mp,R}(t)\Big).\end{aligned}$$

Proof. For x with $\|x - m_\pm(t)\| \leqslant \varepsilon^{4\gamma}$ and y satisfying $\|y\| \leqslant \varepsilon^{-1-\gamma}$ the inequality

$$\|x + \varepsilon g(x)y - m_\pm(t) - \varepsilon g(m_\pm(t))y\| \leqslant \varepsilon^{2\gamma}$$

follows from the Lipschitz continuity of g that results from the boundedness of the derivatives (assumption (N2)). Assume $B \in \mathcal{B}(\mathbb{R}^d)$. The inequality above implies the inclusions

$$\left\{y \in \mathbb{R}^d : x + \varepsilon g(x)y \in B, \|y\| \leqslant \varepsilon^{-1-\gamma}\right\} \subseteq \left\{y \in \mathbb{R}^d : m_\pm(t) + \varepsilon m_\pm(t)y \in B^{+\varepsilon^{2\gamma}}\right\}, \tag{3.9}$$

$$\begin{aligned}\left\{y \in \mathbb{R}^d : x + \varepsilon g(x)y \in B, \|y\| \leqslant \varepsilon^{-1-\gamma}\right\} \supseteq &\left\{y \in \mathbb{R}^d : m_\pm(t) + \varepsilon m_\pm(t)y \in B^{-\varepsilon^{2\gamma}}\right\} \\ &\cap \left\{\|y\| \leqslant \varepsilon^{-1-\gamma}\right\}\end{aligned} \tag{3.10}$$

for $\|x - m_\pm(t)\| \leqslant \varepsilon^{4\gamma}$. Inclusion (3.9) and the following inclusions

$$\begin{aligned}(\Omega_{\mp,R}(t))^{+\varepsilon^{2\gamma}} &\subseteq \Omega_{\mp,R}(t) \cup \Gamma_R^{\varepsilon^\gamma}(t) \cup \left(O_R^{+\varepsilon^{2\gamma}} \backslash O_R\right), \\ ((\Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t) \cap O_R^{-\varepsilon^{2\gamma}})^c)^{+\varepsilon^{2\gamma}} &\subseteq \Omega_{\mp,R}(t) \cup \Gamma_R^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t) \cup \left(O_R \backslash O_R^{-2\varepsilon^{2\gamma}}\right) \cup O_R^c,\end{aligned}$$

verify assertion (i) and (iv). Due to inclusion (3.10) and the inclusions

$$\begin{aligned}((\Omega_{\pm,R}^{\varepsilon^\gamma}(t))^c)^{-\varepsilon^{2\gamma}} &\supseteq \Omega_{\mp,R}(t), \\ (\Omega_{\mp,R}(t))^{-\varepsilon^{2\gamma}} &\supseteq \Omega_{\mp,R}(t) \backslash \left(\Gamma_R^{\varepsilon^\gamma}(t) \cup \left(O_R \backslash O_R^{-\varepsilon^{2\gamma}}\right)\right),\end{aligned}$$

the remaining assertions (ii) and (iii) can be proven. □

3.4.2 Main result

Without any further assumptions except those presented in Section 3.2 we can eventually prove a cumbersome estimate of the transition probability.

Proposition 3.36. *Suppose $R > R^*$ and $\delta, \bar{\delta} > 0$. Define $I := \left[\frac{1}{2} - \frac{\delta}{2}, \frac{1}{2} + \frac{\delta}{2}\right]$. There are constants $\varepsilon_0, \gamma_0 > 0$ such that the following is true*

$$\begin{aligned}&\mathbb{P}_x\left(\hat{\tau}_\pm^\varepsilon \in \left[\frac{1}{2} - \frac{\delta}{2}, \frac{1}{2} + \frac{\delta}{2}\right], \hat{X}_{\hat{\tau}_\pm^\varepsilon}^\varepsilon(x) \in \Omega_{\mp,R}(2T\hat{\tau}_\pm^\varepsilon)\right) \in \\ &\left[-\bar{\delta} + \int_I 2T\left(\nu\left(\frac{1}{\varepsilon}D_{\mp,R}(t)\right) - \nu\left(\frac{1}{\varepsilon}G_{\mp,R}^\varepsilon(t)\right)\right)\right. \\ &\qquad \exp\left\{-2T\int_0^t \left(\nu\left(\frac{1}{\varepsilon}D_{\mp,R}(r)\right) + \nu\left(\frac{1}{\varepsilon}G_{\mp,R}^\varepsilon(r)\right) + \nu\left(\frac{1}{\varepsilon}H_{\mp,R}^\varepsilon(r)\right)\right)\mathrm{d}r\right\}\mathrm{d}t, \\ &\quad \left.\bar{\delta} + \int_I 2T\left(\nu\left(\frac{1}{\varepsilon}D_{\mp,R}(t)\right) + \nu\left(\frac{1}{\varepsilon}G_{\mp,R}^\varepsilon(t)\right)\right)\exp\left\{-2T\int_0^t \nu\left(\frac{1}{\varepsilon}D_{\mp,R}(r)\right)\mathrm{d}r\right\}\mathrm{d}t\right],\end{aligned}$$

for all $x \in \Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(0) \cap O_R^{-\varepsilon^{2\gamma}}$, $\varepsilon \in (0, \varepsilon_0)$ and $\gamma \in (0, \gamma_0)$.

The supposed upper and lower bounds of the transition probability can be altered to be independent of ε and R through bringing the limit measure μ into play and by posing assumptions on the error terms $\nu\left(\frac{1}{\varepsilon}G_{\mp,R}^\varepsilon(t)\right)$ and $\nu\left(\frac{1}{\varepsilon}H_{\mp,R}^\varepsilon(t)\right)$ and the main term $\nu\left(\frac{1}{\varepsilon}D_{\mp,R}(t)\right)$. At first some definitions follow.

Definition 3.37. *For $s \geqslant 0$ consider the homogeneous equation $\frac{\mathrm{d}}{\mathrm{d}t}\tilde{x}(t) = -\nabla U(\tilde{x}(t), s)$, $t \geqslant 0$ and let $\widetilde{\Omega}_{\pm}(s)$ be the basin of attraction of $m_{\pm}(s)$. Let $\widetilde{\Omega}_{\pm,R}(s) := \widetilde{\Omega}_{\pm}(s) \cap O_R$ and define*

$$\widetilde{D}_{\mp}(t) := \left\{ y \in \mathbb{R}^d : m_{\pm}(t) + g(m_{\pm}(t))y \in \widetilde{\Omega}_{\mp}(t) \right\},$$
$$\widetilde{D}_{\mp,R}(t) := \left\{ y \in \mathbb{R}^d : m_{\pm}(t) + g(m_{\pm}(t))y \in \widetilde{\Omega}_{\mp,R}(t) \right\}.$$

Due to assumption (T) the period length $2T$ equals to $\frac{c_{per}}{\varepsilon^{\alpha} l(\varepsilon^{-1})}$. Introduce the succeeding conditions for $R \geqslant R^*$.

(H1) Demand that $2T\nu\left(\frac{1}{\varepsilon}G^{\varepsilon}_{\mp,R}(t)\right)$ tends to zero uniformly in t as ε converges to zero and the convergence of $2T\nu\left(\frac{1}{\varepsilon}H^{\varepsilon}_{\mp,R}(t)\right)$ to $c_{per}\mu(H_{\mp,R}(t))$ is uniform in time as $\varepsilon \to 0$.

(H2) The difference $2T\left(\nu\left(\frac{1}{\varepsilon}D_{\mp,R}(t)\right) - \nu\left(\frac{1}{\varepsilon}\widetilde{D}_{\mp,R}(t)\right)\right)$ is a null sequence as ε tends to zero while the convergence is uniform in t. Additionally the convergence of $2T\nu\left(\frac{1}{\varepsilon}\widetilde{D}_{\mp,R}(t)\right)$ to $c_{per}\mu\left(\widetilde{D}_{\mp,R}(t)\right)$ as $\varepsilon \to 0$ given by the regular variation property shall be uniform in t.

Under these hypotheses the following consequence of the previous proposition is immediate.

Theorem 3.38. *Let* (H1) *and* (H2) *be satisfied. For all $\delta, \bar{\delta} > 0$ there exist $\gamma, \varepsilon_0 > 0$ and an $R > R^*$ such that for all $\varepsilon < \varepsilon_0$ the assertion*

$$\left| \mathbb{P}_x\left(\hat{\tau}^{\varepsilon}_{\pm} \in \left[\frac{1}{2} - \frac{\delta}{2}, \frac{1}{2} + \frac{\delta}{2}\right], \hat{X}^{\varepsilon}_{\hat{\tau}^{\varepsilon}_{\pm}}(x) \in \Omega_{\mp,R}(2T\hat{\tau}^{\varepsilon}_{\pm}) \right) - I^{\delta}_{\pm} \right| \leqslant \bar{\delta}$$

holds uniformly in $x \in \Omega^{\varepsilon^{\gamma},\varepsilon^{2\gamma}}_{\pm}(0) \cap O_R^{-\varepsilon^{2\gamma}}$ with

$$I^{\delta}_{\pm} := \exp\left(-\int_0^{(1-\delta)/2} c_{per}\mu\left(\widetilde{D}_{\mp}(t)\right)\mathrm{d}t \right) - \exp\left(-\int_0^{(1+\delta)/2} c_{per}\mu\left(\widetilde{D}_{\mp}(t)\right)\mathrm{d}t \right).$$

Remark 3.39. In Subsection 3.1.2 among others we calculated the probability that the two-valued Markov chain C jumps from -1 to 1 within the time interval $[T(1-\delta), T(1+\delta)]$. This probability given in equation (3.5) for $k = 0$ shows the same structure as the term $I^{\delta}_{\pm}$ that satisfies

$$I^{\delta}_{\pm} = \delta\, c_{per}\mu\left(\widetilde{D}_{\mp}\left(\frac{1}{2}\right)\right) \mathrm{e}^{-\int_0^{1/2} c_{per}\mu\left(\widetilde{D}_{\mp}(t)\right)\mathrm{d}t} + \mathrm{O}(\delta^2) =: \delta f(c_{per}) + \mathrm{O}(\delta^2).$$

The funtion f attains its maximum at

$$c_{per} = \left(\int_0^{1/2} \mu\left(\widetilde{D}_{\mp}(t)\right) \mathrm{d}t \right)^{-1}.$$

It is optimal to choose

$$2T(\varepsilon) = \left(\varepsilon^{\alpha} l(\varepsilon^{-1}) \int_0^{1/2} \mu\left(\widetilde{D}_{\mp}(t)\right) \mathrm{d}t \right)^{-1},$$

which is similar to the optimal period obtained for the Markov chain in Proposition 3.9 ($k = 0$).

We owe the proof of Proposition 3.36.

Proof. **Upper bound**

Let N_t be the counting process, associated with the compensated Poisson process η, with $N_t := \sum_{s\leqslant t} \mathbf{1}_{\{\Delta\eta_s\neq 0\}}$, which is a Poisson process with parameter $\beta^\varepsilon = \nu(\{\|x\| \geqslant \varepsilon^{-\rho}\})$. Choose $\rho \in \left(\frac{2}{3}, 1\right)$ and $k_\varepsilon := \lfloor \varepsilon^{-\kappa} \rfloor$ for any $\kappa \in \left[\alpha(1-\rho), \frac{\alpha\rho}{2}\right]$. Then for $x \in \Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(0) \cap O_R^{-\varepsilon^{2\gamma}}$ we obtain

$$\begin{aligned}&\mathbb{P}_x\left(\hat{\tau}_\pm^\varepsilon \in I, \hat{X}^\varepsilon_{\hat{\tau}^\varepsilon_\pm} \in \Omega_{\mp,R}(2T\hat{\tau}_\pm^\varepsilon)\right)\\ &\leqslant \mathbb{P}\left(N_{T(1+\delta)} > k_\varepsilon\right) + \mathbb{P}_x\left(N_{T(1+\delta)} \leqslant k_\varepsilon, \hat{\tau}_\pm^\varepsilon \in I, \hat{X}^\varepsilon_{\hat{\tau}^\varepsilon_\pm} \in \Omega_{\mp,R}(2T\hat{\tau}_\pm^\varepsilon)\right).\end{aligned}$$

Due to the Poisson distribution of N_t the probability that more than k_ε jumps of η occur before $2T\left(\frac{1}{2}+\frac{\delta}{2}\right) = T(1+\delta)$ can be estimated through Lemma 3.33 as

$$\mathbb{P}\left(N_{T(1+\delta)} > k_\varepsilon\right) = \sum_{k=k_\varepsilon+1}^{\infty} \mathrm{e}^{-T\beta^\varepsilon(1+\delta)} \frac{(T\beta^\varepsilon\,(1+\delta))^k}{k!} \leqslant \frac{(T\beta^\varepsilon\,(1+\delta))^{k_\varepsilon+1}}{(k_\varepsilon+1)!} \leqslant \left(CT\beta^\varepsilon k_\varepsilon^{-1}\right)^{k_\varepsilon+1}$$

for some $C > 0$, where the last inequality follows from the formula of Stirling that guarantees $n! \geqslant \sqrt{2\pi n}\, n^n \mathrm{e}^{-n}$ for $n \in \mathbb{N}$. Because of

$$\mathbb{P}\left(N_{T(1+\delta)} > k_\varepsilon\right) \leqslant \left(CT\beta^\varepsilon k_\varepsilon^{-1}\right)^{k_\varepsilon+1} \leqslant \left(C\frac{c_{per}}{2\varepsilon^\alpha l(\varepsilon^{-1})}\,\varepsilon^{\alpha\rho} l(\varepsilon^{-\rho})\,\frac{1}{\lfloor \varepsilon^{-\kappa}\rfloor}\right)^{\lfloor \varepsilon^{-\kappa}\rfloor+1}$$

the upper bound is a null sequence as ε tends to zero if $\kappa > \alpha(1-\rho)$. Let us study the dynamics under the condition $\left\{N_{T(1+\delta)} \leqslant k_\varepsilon\right\}$. As usual, we distinguish between exits from a well between the normalized jump times $\hat{\tau}_k = \frac{\tau_k}{2T}$ of η and those occuring exactly at $\hat{\tau}_k$. For $x \in \Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(0) \cap O_R^{-\varepsilon^{2\gamma}}$ we estimate

$$\begin{aligned}&\mathbb{P}_x\left(N_{T(1+\delta)} \leqslant k_\varepsilon, \hat{X}^\varepsilon_{\hat{\tau}^\varepsilon_\pm}(x) \in \Omega_{\mp,R}(2T\hat{\tau}_\pm^\varepsilon)\right)\\ &\leqslant \sum_{k=1}^{k_\varepsilon}\Big[\mathbb{P}_x\left(\hat{\tau}_\pm^\varepsilon = \hat{\tau}_k \in I, \hat{X}^\varepsilon_{\hat{\tau}_k}(x) \in \Omega_{\mp,R}(2T\hat{\tau}_k)\right)\\ &\qquad + \mathbb{P}_x\left(\hat{\tau}_\pm^\varepsilon \in (\hat{\tau}_{k-1},\hat{\tau}_k)\cap I, \hat{X}^\varepsilon_{\hat{\tau}^\varepsilon_\pm}(x) \in \Omega_{\mp,R}(2T\hat{\tau}_\pm^\varepsilon)\right)\Big].\end{aligned}$$

A direct transition from one bounded and reduced domain to the other one is impossible on the event $E_{\hat{\tau}_{k-1},\hat{\tau}_k}(\hat{X}^\varepsilon_{\hat{\tau}_{k-1}})$. This implies

$$\begin{aligned}&\mathbb{P}_x\left(\hat{\tau}_\pm^\varepsilon \in (\hat{\tau}_{k-1},\hat{\tau}_k)\cap I, \hat{X}^\varepsilon_{\hat{\tau}^\varepsilon_\pm}(x) \in \Omega_{\mp,R}(2T\hat{\tau}_\pm^\varepsilon)\right)\\ &= \mathbb{E}_x\left(\prod_{j=0}^{k-2}\mathbf{1}\left(B^{\pm,j}_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x))\right)\mathbf{1}\left(F^{\pm}_{\hat{\tau}_{k-1},\hat{\tau}_k}(\hat{X}^\varepsilon_{\hat{\tau}_{k-1}})\right)\right)\\ &\leqslant \mathbb{E}_x\left(\prod_{j=0}^{k-2}\mathbf{1}\left(B^{\pm,j}_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x))\right)\mathbb{E}\left(\mathbf{1}\left(E^c_{\hat{\tau}_{k-1},\hat{\tau}_k}(\hat{X}^\varepsilon_{\hat{\tau}_{k-1}}(x))\right)\Big|\mathcal{F}_{\tau_{k-1}}\right)\right).\end{aligned}$$

Due to the strong Markov property of the homogenised process $(\hat{X}^\varepsilon_t, t)_{t\geqslant 0}$ the involved conditional expectation is a function h depending on the initial value $\hat{X}^\varepsilon_{\hat{\tau}_{k-1}}$ and the starting time $\hat{\tau}_{k-1}$. It holds

$$\begin{aligned}
&\mathbb{P}_x\left(\hat{\tau}^\varepsilon_\pm \in (\hat{\tau}_{k-1}, \hat{\tau}_k) \cap I, \hat{X}^\varepsilon_{\hat{\tau}^\varepsilon_\pm}(x) \in \Omega_{\mp,R}(2T\hat{\tau}^\varepsilon_\pm)\right)\\
&\leqslant \mathbb{E}_x\left(\prod_{j=0}^{k-2} \mathbf{1}\left(B^{\pm,j}_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x))\right) \sup_{x_0\in\Omega^{\varepsilon^\gamma}_{\pm,R}(2T\hat{\tau}_{k-1})} h(x_0, \hat{\tau}_{k-1})\right)\\
&\leqslant \sup_{r\geqslant 0} \sup_{x_0\in\Omega^{\varepsilon^\gamma}_{\pm,R}(r)} \mathbb{P}_x\left(\sup_{t\leqslant\tau_1} \left\|X^\xi_{r,t}(x_0) - X^0_{r,t}(x_0)\right\| \geqslant \varepsilon^{4\gamma}\right).
\end{aligned}$$

From Theorem 3.18 we derive the upper bound $\mathrm{e}^{-\varepsilon^{-p}}$ for the probability of an exit between two normalized jump times of η.

It is left to treat exits arising from big jumps. Remember Definition 3.28 to get

$$\begin{aligned}
&\mathbb{P}_x\left(\hat{\tau}^\varepsilon_\pm = \hat{\tau}_k \in I, \hat{X}^\varepsilon_{\hat{\tau}_k}(x) \in \Omega_{\mp,R}(2T\hat{\tau}_k)\right)\\
&= \mathbb{E}_x\left(\mathbf{1}_{\{\hat{\tau}_k\in I\}} \prod_{j=0}^{k-2} \mathbf{1}\left(B^{\pm,j}_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x))\right) \mathbf{1}\left(C^{\pm,k-1}_{\hat{\tau}_{k-1},\hat{\tau}_k}(\hat{X}^\varepsilon_{\hat{\tau}_{k-1}}(x))\right)\right)\\
&\leqslant \int_I \int_0^{t_k} \cdots \int_0^{t_2} \Bigg[(2T\beta^\varepsilon)^k\, \mathrm{e}^{-2T\beta^\varepsilon t_k} \prod_{j=0}^{k-2} \sup_{x\in\Omega^{\varepsilon^\gamma}_{\pm,R}(2Tt_j)} \mathbb{P}\left(B^{\pm,j}_{t_j,t_{j+1}}(x)\right)\\
&\qquad \sup_{x\in\Omega^{\varepsilon^\gamma}_{\pm,R}(2Tt_{k-1})} \mathbb{P}\left(C^{\pm,k-1}_{t_{k-1},t_k}(x)\right)\Bigg] \mathrm{d}t_1 \ldots \mathrm{d}t_k,
\end{aligned}$$

with $t_0 = 0$ while the inequality follows from Lemma 3.30. Multiply the integrand with the identity

$$1 = \prod_{j=0}^{k-1} \mathbf{1}_{\{2T(t_{j+1}-t_j)\geqslant R_{\varepsilon^\gamma}\}} + 1 - \prod_{j=0}^{k-1} \mathbf{1}_{\{2T(t_{j+1}-t_j)\geqslant R_{\varepsilon^\gamma}\}}$$

and exploit Lemma 3.31 (i) and (ii) for the main part of the probability of a transition at $\hat{\tau}_k$ to end up with

$$\begin{aligned}
&\mathbb{P}_x(\hat{\tau}^\varepsilon_\pm = \hat{\tau}_k \in I, \hat{X}^\varepsilon_{\hat{\tau}_k}(x) \in \Omega_{\mp,R}(2T\hat{\tau}_k))\\
&\leqslant \int_I \int_0^{t_k} \cdots \int_0^{t_2} \Bigg[(2T)^k \mathrm{e}^{-2T\beta^\varepsilon t_k} \prod_{j=1}^{k-1} \left(\beta^\varepsilon \mathrm{e}^{-\varepsilon^{-p}} + \sup_{x:\,\|x-m_\pm(t_j)\|\leqslant\varepsilon^{4\gamma}} \nu\left(B^c_{\varepsilon^{-\rho}}(0) \cap \frac{1}{\varepsilon} D^{\varepsilon^\gamma}_{\pm,R}(x,t_j)\right)\right)\\
&\quad \left(\beta^\varepsilon \mathrm{e}^{-\varepsilon^{-p}} + \sup_{x:\,\|x-m_\pm(t_k)\|\leqslant\varepsilon^{4\gamma}} \nu\left(\frac{1}{\varepsilon} D_{\mp,R}(x,t_k)\right)\right)\Bigg] \mathrm{d}t_1 \ldots \mathrm{d}t_k\\
&\quad + \int_I \int_0^{t_k} \cdots \int_0^{t_2} (2T\beta^\varepsilon)^k \mathrm{e}^{-2T\beta^\varepsilon t_k} \left(1 - \prod_{j=0}^{k-1} \mathbf{1}_{\{2T(t_{j+1}-t_j)\geqslant R_{\varepsilon^\gamma}\}}\right) \mathrm{d}t_1 \ldots \mathrm{d}t_k.
\end{aligned}$$

The error term consisting of the last iterated integral of the upper bound above can be treated with Lemma 3.32 (i) with $c := \frac{R_{\varepsilon^\gamma}}{2T}$, $C := t_k$ and $n := k-1$ as presented in the sequel:

$$\begin{aligned}
&\int_I \int_0^{t_k} \dots \int_0^{t_2} (2T\beta^\varepsilon)^k \,\mathrm{e}^{-2T\beta^\varepsilon t_k} \left(1 - \prod_{j=0}^{k-1} \mathbf{1}_{\{2T(t_{j+1}-t_j) \geqslant R_{\varepsilon^\gamma}\}} \right) \mathrm{d}t_1 \dots \mathrm{d}t_k \\
&\leqslant \int_I \mathrm{e}^{-2T\beta^\varepsilon t} (2T\beta^\varepsilon)^k \frac{t^{k-1}}{(k-1)!} \left(1 - \left(1 - \frac{kR_{\varepsilon^\gamma}}{2Tt} \right)^{k-1} \right) \mathrm{d}t.
\end{aligned}$$

For integers $k \leqslant k_\varepsilon$ and $t \in I$ the term $\frac{k(k-1)R_{\varepsilon^\gamma}}{2Tt}$ is at most of order $\varepsilon^{\alpha - 2\kappa}$ and thus converges to zero as $\varepsilon \to 0$ and $\kappa < \frac{\alpha}{2}$ which holds due to $\kappa < \frac{\rho\alpha}{2}$. This and the series representation $\log(1+x) = \sum_{j=1}^\infty \frac{(-1)^{j+1}x^j}{j}$ for $|x| < 1$ enable us to estimate

$$\begin{aligned}
\left(1 - \frac{R_{\varepsilon^\gamma} k}{2Tt} \right)^{k-1} &= \exp\left((k-1) \left(-\frac{R_{\varepsilon^\gamma} k}{2Tt} - \frac{(R_{\varepsilon^\gamma} k)^2}{2(2Tt)^2} - \frac{(R_{\varepsilon^\gamma} k)^3}{3(2Tt)^3} - \dots \right) \right) \\
&\geqslant \exp\left(-\frac{R_{\varepsilon^\gamma} k(k-1)}{2Tt} - \frac{(R_{\varepsilon^\gamma}(k-1))^2}{2(2Tt)^2} - \frac{(R_{\varepsilon^\gamma} k(k-1))^3}{3(2Tt)^3} - \dots \right) \\
&= 1 - \frac{R_{\varepsilon^\gamma} k(k-1)}{2Tt}
\end{aligned}$$

and these considerations induce

$$\begin{aligned}
&\int_I \int_0^{t_k} \dots \int_0^{t_2} (2T\beta^\varepsilon)^k \,\mathrm{e}^{-2T\beta^\varepsilon t_k} \left(1 - \prod_{j=0}^{k-1} \mathbf{1}_{\{2T(t_{j+1}-t_j) \geqslant R_{\varepsilon^\gamma}\}} \right) \mathrm{d}t_1 \dots \mathrm{d}t_k \\
&\leqslant \int_I (2T\beta^\varepsilon)^k \,\mathrm{e}^{-2T\beta^\varepsilon t} \frac{t^{k-1}}{(k-1)!} \frac{R_{\varepsilon^\gamma}}{2T} \frac{k(k-1)}{t} \,\mathrm{d}t \\
&\leqslant k_\varepsilon \beta^\varepsilon R_{\varepsilon^\gamma}.
\end{aligned}$$

Adding up these errors for $k \in \{1, \dots, k_\varepsilon\}$ yields an error of order $\varepsilon^{\alpha\rho - 2\kappa}$ that vanishes as $\varepsilon \to 0$ because $\kappa < \frac{\alpha\rho}{2}$. Because of the well known series representation of e^x and Lemma 3.32 (ii) the probability for a transition to the other bounded and reduced domain through a big jump within the time interval I can be bounded by

$$\begin{aligned}
&\sum_{k=1}^{k_\varepsilon} \mathbb{P}_x \left(\hat{\tau}_\pm^\varepsilon = \hat{\tau}_k \in I, \hat{X}_{\hat{\tau}_k}^\varepsilon(x) \in \Omega_{\mp,R}(2T\hat{\tau}_k) \right) \\
&\leqslant k_\varepsilon^2 \beta^\varepsilon R_{\varepsilon^\gamma} + \int_I \Bigg[2T\,\mathrm{e}^{-2T\beta^\varepsilon t} \Big(\beta^\varepsilon \mathrm{e}^{-\varepsilon^{-p}} + \sup_{x\colon \|x - m_\pm(t)\| \leqslant \varepsilon^{4\gamma}} \nu\Big(\frac{1}{\varepsilon} D_{\mp,R}(x,t) \Big) \Big) \\
&\quad \sum_{k=1}^\infty \Bigg(\int_0^t \int_0^{t_{k-1}} \dots \int_0^{t_2} \Bigg\{ \prod_{j=1}^{k-1} \Big(\beta^\varepsilon \mathrm{e}^{-\varepsilon^{-p}} + \sup_{x\colon \|x - m_\pm(t_j)\| \leqslant \varepsilon^{4\gamma}} \nu\Big(B_{\varepsilon^{-\rho}}^c(0) \cap \frac{1}{\varepsilon} D_{\pm,R}^{\varepsilon^\gamma}(x,t_j) \Big) \Big) \\
&\quad (2T)^{k-1} \Bigg\} \mathrm{d}t_1 \dots \mathrm{d}t_{k-1} \Bigg) \Bigg] \mathrm{d}t \\
&\leqslant k_\varepsilon^2 \beta^\varepsilon R_{\varepsilon^\gamma} + \int_I \Bigg[2T \Big(\beta^\varepsilon \mathrm{e}^{-\varepsilon^{-p}} + \sup_{x\colon \|x - m_\pm(t)\| \leqslant \varepsilon^{4\gamma}} \nu\Big(\frac{1}{\varepsilon} D_{\mp,R}(x,t) \Big) \Big) \\
&\quad \exp\left(-2T \int_0^t \Big(\inf_{x\colon \|x - m_\pm(r)\| \leqslant \varepsilon^{4\gamma}} \nu\Big(\Big(\frac{1}{\varepsilon} D_{\pm,R}^{\varepsilon^\gamma}(x,r) \Big)^c \Big) - \beta^\varepsilon \mathrm{e}^{-\varepsilon^{-p}} \Big) \mathrm{d}r \right) \Bigg] \mathrm{d}t.
\end{aligned}$$

The exponential terms in ε only produce exponentially small errors. From Lemma 3.35 (i) and (ii), $2T\nu(B^c_{\varepsilon^{-1-\gamma}}(0)) \approx \varepsilon^{\alpha\gamma}$ and all previous arguments the supposed upper bound follows.

Lower bound

Again choose $\rho \in \left(\frac{2}{3}, 1\right)$ and $k_\varepsilon = \lfloor \varepsilon^{-\kappa} \rfloor$ for $\kappa \in [\alpha(1-\rho), \frac{\alpha\rho}{2}]$. Exits arising from small jumps or very late big jumps can immediately be disregarded as done in the first estimate below

$$
\begin{aligned}
&\mathbb{P}_x\left(\hat{\tau}^\varepsilon_\pm \in I, \hat{X}^\varepsilon_{\hat{\tau}^\varepsilon_\pm}(x) \in \Omega_{\mp,R}(2T\hat{\tau}^\varepsilon_\pm)\right) \\
&\geqslant \sum_{k=1}^{k_\varepsilon} \mathbb{P}_x\left(\hat{\tau}^\varepsilon_\pm = \hat{\tau}_k \in I, \hat{X}^\varepsilon_{\hat{\tau}_k}(x) \in \Omega_{\mp,R}(2T\hat{\tau}_k)\right) \\
&\geqslant \sum_{k=1}^{k_\varepsilon} \mathbb{E}_x\left(\mathbf{1}_{\{\hat{\tau}_k \in I\}} \prod_{j=0}^{k-2} \mathbf{1}\left(\bar{B}^{\pm,j}_{\hat{\tau}_j,\hat{\tau}_{j+1}}(\hat{X}^\varepsilon_{\hat{\tau}_j}(x))\right) \mathbf{1}\left(C^{\pm,k-1}_{\hat{\tau}_{k-1},\hat{\tau}_k}(\hat{X}^\varepsilon_{\hat{\tau}_{k-1}}(x))\right)\right) \\
&\geqslant \sum_{k=1}^{k_\varepsilon} \int_I \int_0^{t_k} \cdots \int_0^{t_2} \Bigg[(2T\beta^\varepsilon)^k \mathrm{e}^{-2T\beta^\varepsilon t_k} \prod_{j=0}^{k-2} \inf_{x \in \Omega^{\varepsilon^\gamma}_{\pm,R}(2Tt_j)} \mathbb{P}\left(\bar{B}^{\pm,j}_{t_j,t_{j+1}}(x)\right) \\
&\qquad \inf_{x \in \Omega^{\varepsilon^\gamma}_{\pm,R}(2Tt_{k-1})} \mathbb{P}\left(C^{\pm,k-1}_{t_{k-1},t_k}(x)\right)\Bigg] \mathrm{d}t_1 \ldots \mathrm{d}t_k,
\end{aligned}
$$

while the last estimate is a consequence of Lemma 3.30. Introducing the factor $\prod_{j=0}^{k-1} \mathbf{1}_{\{2T(t_{j+1}-t_j) \geqslant R_{\varepsilon^\gamma}\}} \in [0,1]$ and applying Lemma 3.31 (iii) and (iv) yields

$$
\begin{aligned}
&\mathbb{P}_x\left(\hat{\tau}^\varepsilon_\pm \in I, \hat{X}^\varepsilon_{\hat{\tau}^\varepsilon_\pm}(x) \in \Omega_{\mp,R}(2T\hat{\tau}^\varepsilon_\pm)\right) \\
&\geqslant \sum_{k=1}^{k_\varepsilon} \int_I \int_0^{t_k} \cdots \int_0^{t_2} \Bigg[(2T)^k \, \mathrm{e}^{-2T\beta^\varepsilon t_k} \prod_{j=1}^{k-1} \inf_{x:\, \|x - m_\pm(t_j)\| \leqslant \varepsilon^\gamma} \nu\left(B^c_{\varepsilon^{-\rho}}(0) \cap \frac{1}{\varepsilon} D^{\varepsilon^\gamma,\varepsilon^{2\gamma}}_{\pm,R}(x,t_j)\right) \\
&\quad \inf_{x:\, \|x - m_\pm(t_k)\| \leqslant \varepsilon^\gamma} \nu\left(\frac{1}{\varepsilon} D_{\mp,R}(x,t_k)\right) \left(1 - \mathrm{e}^{-\varepsilon^{-p}}\right)^k \prod_{j=0}^{k-1} \mathbf{1}_{\{2T(t_{j+1}-t_j) \geqslant R_{\varepsilon^\gamma}\}}\Bigg] \mathrm{d}t_1 \ldots \mathrm{d}t_k.
\end{aligned}
$$

Omitting $\prod_{j=0}^{k-1} \mathbf{1}_{\{2T(t_{j+1}-t_j) \geqslant R_{\varepsilon^\gamma}\}}$ again produces an error of order $\varepsilon^{\alpha\rho - 2\kappa}$ which is a null sequence under our assumptions. An extension of the finite sum for $k \in \{1, \ldots, k_\varepsilon\}$ to an infinite series creates another error which is smaller than $\mathbb{P}(N_{T(1+\delta)} > k_\varepsilon)$ and thus a null sequence as ε tends to zero if $\kappa > \alpha(1-\rho)$. Again ignore the terms $\mathrm{e}^{-\varepsilon^{-p}}$, use Lemma 3.35 (iii) and (iv) and $2T\nu(B^c_{\varepsilon^{-1-\gamma}}) \approx \varepsilon^{\alpha\gamma}$ to complete the proof. □

3.4.3 The transition probability for the Marcus stochastic differential equation

The handling of Z^ε solving (3.2) is usually quite laborious, but again we can adopt the methods from the congeneric Itô case. The main distinction concerning transitions from $\Omega_\pm(2Tt)$ to $\Omega_\mp(2Tt)$ is the importance of big jumps W_1 with $\varphi(1; m_\pm(t), \varepsilon W_1) = \varepsilon\varphi(1; m_\pm(t), W_1) \in \Omega_\mp(2Tt)$ where

$$
\frac{\mathrm{d}}{\mathrm{d}t}\varphi(t; x, y) = g(\varphi(t; x, y))y, \quad \varphi(0; x, y) = x,
$$

instead of those with $m_\pm(t) + \varepsilon g(m_\pm(t))W_1 \in \Omega_\mp(2Tt)$.

Definition 3.40. *Define the scaled process $\hat{Z}_t^\varepsilon = Z_{2Tt}^\varepsilon$, the exit time*

$$\hat{\tau}_\pm^{M,\varepsilon} = \inf\left\{t \geqslant 0 : \hat{Z}_t^\varepsilon \notin \Omega_{\pm,R}^{\varepsilon^\gamma}(2Tt)\right\},$$

and the relevant jump sets

$$\begin{aligned}
D_{\mp,R}^M(t) &= \left\{y \in \mathbb{R}^d : \varphi(1; m_\pm(t), y) \in \Omega_\mp(2Tt)\right\},\\
G_{\mp,R}^{M,\varepsilon}(t) &= \left\{y \in \mathbb{R}^d : \varphi(1; m_\pm(t), y) \in \Gamma_R^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(2Tt) \cup \left(O_R^{+\varepsilon^{2\gamma}} \backslash O_R^{-2\varepsilon^{2\gamma}}\right)\right\},\\
H_{\mp,R}^{M,\varepsilon}(t) &= \left\{y \in \mathbb{R}^d : \varphi(1; m_\pm(t), y) \in \left(O_R^{-2\varepsilon^{2\gamma}}\right)^c\right\},\\
H_{\mp,R}^M(t) &= \left\{y \in \mathbb{R}^d : \varphi(1; m_\pm(t), y) \in O_R^c\right\}.
\end{aligned}$$

Proposition 3.41. *Suppose $R > R^*$ and $\delta, \bar{\delta} > 0$. Define $I = \left[\frac{1}{2} - \frac{\delta}{2}, \frac{1}{2} + \frac{\delta}{2}\right]$. There are constants $\varepsilon_0, \gamma_0 > 0$ such that the following is true*

$$\begin{aligned}
&\mathbb{P}_x\left(\hat{\tau}_\pm^{M,\varepsilon} \in \left[\frac{1}{2} - \frac{\delta}{2}, \frac{1}{2} + \frac{\delta}{2}\right], \hat{Z}_{\hat{\tau}_\pm^{M,\varepsilon}}^\varepsilon(x) \in \Omega_{\mp,R}(2T\hat{\tau}_\pm^{M,\varepsilon})\right) \in\\
&\Bigg[-\bar{\delta} + \int_I 2T\left(\nu\left(\frac{1}{\varepsilon}D_{\mp,R}^M(t)\right) - \nu\left(\frac{1}{\varepsilon}G_{\mp,R}^{M,\varepsilon}(t)\right) - \nu\left(B_{\varepsilon^{-1}R}^c(0)\right)\right)\\
&\quad\exp\left\{-2T\int_0^t \left(\nu\left(\frac{1}{\varepsilon}D_{\mp,R}^M(r)\right) + \nu\left(\frac{1}{\varepsilon}G_{\mp,R}^{M,\varepsilon}(r)\right) + \nu\left(\frac{1}{\varepsilon}H_{\mp,R}^{M,\varepsilon}(r)\right) + \nu\left(B_{\varepsilon^{-1}R}^c(0)\right)\right)\mathrm{d}r\right\}\mathrm{d}t,\\
&\quad\bar{\delta} + \int_I 2T\left(\nu\left(\frac{1}{\varepsilon}D_{\mp,R}^M(t)\right) + \nu\left(\frac{1}{\varepsilon}G_{\mp,R}^{M,\varepsilon}(t)\right) + \nu\left(B_{\varepsilon^{-1}R}^c(0)\right)\right)\\
&\quad\quad\exp\left\{-2T\int_0^t \left(\nu\left(\frac{1}{\varepsilon}D_{\mp,R}^M(r)\right) - \nu\left(B_{\varepsilon^{-1}R}^c(0)\right)\right)\mathrm{d}r\right\}\mathrm{d}t\Bigg],
\end{aligned}$$

for all $x \in \Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(0) \cap O_R^{-\varepsilon^{2\gamma}}$, $\varepsilon \in (0, \varepsilon_0)$ and $\gamma \in (0, \gamma_0)$.

Proof. The proof of Proposition 3.36 can be retraced. A minor difference occurs in the preparatory Lemma 3.35 which explains the slightly different error terms in the upper and lower bound. If $\|x - z\|$ is small, $\varphi(1; x, y)$ belongs to the neighbourhood of $\varphi(1; z, y)$ for y chosen from a bounded set. From the Lipschitz continuity of all components of g with Lipschitz constant C_g, Gronwall's Lemma, εW_1 taken from $B_R(0)$, and $x \in B_{\varepsilon^{4\gamma}}(m_\pm(t))$ we can deduce

$$\begin{aligned}
&\|\varphi(1; x, \varepsilon W_1) - \varphi(1; m_\pm(t), \varepsilon W_1)\|\\
&\leqslant \|x - m_\pm(t)\| + \varepsilon\int_0^1 \|\left(g(\varphi(r; x, \varepsilon W_1)) - g(\varphi(r; m_\pm(t), \varepsilon W_1))\right) W_1\|\,\mathrm{d}r\\
&\leqslant \|x - m_\pm(t)\|\,\mathrm{e}^{C_g R},\\
&\leqslant \varepsilon^{3\gamma}.
\end{aligned}$$

This produces the additional error term $\nu\left(B_{\varepsilon^{-1}R}^c(0)\right)$. □

Definition 3.42. *Analogously to* $\widetilde{D}_{\mp}(t)$ *define*

$$\widetilde{D}^{M}_{\mp}(t) = \left\{ y \in \mathbb{R}^d : \varphi(1; m_{\pm}(t), y) \in \widetilde{\Omega}_{\mp}(t) \right\}.$$

As for the Itô case introduce the hypotheses (H1) and (H2) with $G^{M,\varepsilon}_{\mp,R}(t)$, $H^{M,\varepsilon}_{\mp,R}(t)$, $H^{M}_{\mp,R}(t)$, $D^{M}_{\pm,R}(t)$ and $\widetilde{D}^{M}_{\pm,R}(t)$ instead of their Itô analogues. Call the assumptions (HM1) and (HM2).

Theorem 3.43. *Assume* (HM1) *and* (HM2) *are fulfilled. For all* $\delta, \bar{\delta} > 0$ *there exist* $\gamma, \varepsilon_0 > 0$ *and an* $R > R^*$ *such that for all* $\varepsilon < \varepsilon_0$ *the estimate*

$$\left| \mathbb{P}_x\Big(\hat{\tau}^{M,\varepsilon}_{\pm} \in \Big[\frac{1}{2} - \frac{\delta}{2}, \frac{1}{2} + \frac{\delta}{2}\Big], \hat{Z}^{\varepsilon}_{\hat{\tau}^{M,\varepsilon}_{\pm}}(x) \in \Omega_{\mp,R}\left(2T\hat{\tau}^{M,\varepsilon}_{\pm}\right) \Big) - I^{M,\delta}_{\pm} \right| \leqslant \bar{\delta}$$

holds uniformly for all $x \in \Omega^{\varepsilon^{\gamma},\varepsilon^{2\gamma}}_{\pm}(0) \cap O^{-\varepsilon^{2\gamma}}_{R}$ *with*

$$I^{M,\delta}_{\pm} := \exp\left(- \int_0^{(1-\delta)/2} c_{per}\mu\left(\widetilde{D}^{M}_{\mp}(t)\right) \mathrm{d}t \right) - \exp\left(- \int_0^{(1+\delta)/2} c_{per}\mu\left(\widetilde{D}^{M}_{\mp}(t)\right) \mathrm{d}t \right).$$

The optimal choice of $T(\varepsilon)$ is given by $2T(\varepsilon) = \left(\varepsilon^{\alpha} l(\varepsilon^{-1}) \int_0^{1/2} \mu(\widetilde{D}^{M}_{\mp}(t))\,\mathrm{d}t\right)^{-1}$.

Chapter 4

Metastable behaviour of a jump diffusion driven by a periodic additive process

In the current chapter, we focus our attention on the stochastic differential equation

$$Y_t^\varepsilon(y) = y - \int_0^t \nabla V(Y_s^\varepsilon)\,\mathrm{d}s + \varepsilon A_t^T, \tag{4.1}$$

where the source of periodicity lies within the perturbation process A^T and not in the drift term V. Let V be a double-well potential fulfilling conditions (V1)-(V4) (Section 2.1). Additionally the decomposition (2.4) of the additive process A^T holds true and assumptions (A1) and (A2) are satisfied (Subsection 2.4.1 page 22). In particular, we assume that the local characteristics of the process A^T contain an $\alpha\left(\frac{t}{2T}\right)$-stable Lévy measure where $\alpha\colon \mathbb{R}_+ \to [\alpha_*, \alpha^*]$ with $0 < \alpha_* < \alpha^* < 2$ is 1-periodic and has a unique minimum at $a \in (0, 1)$.
At the beginning a simplification of the essential dynamics of the solution Y^ε is made by considering a time-continuous two-state Markov chain with a periodic Q-matrix and transition rates of order $\varepsilon^{\alpha(t/2T)}$. This Markov chain is almost periodic if the half period T is chosen appropriately according to the noise intensity ε. Afterwards the exit times of the jump diffusion Y^ε from regions of attraction are examined which reveals the same critical time scale as for the approximative Markov chain.

4.1 Approximation by a two-state Markov chain

In the stochastic differential equation (4.1) the source of periodicity is the noise term which is an additive process with time-dependent stability index α. For transitions the lowest value α_* of the 1-periodic function α will be decisive, since then the jump affinity is the greatest. Define the non-autonomous Markov chain $\mathcal{C} = (\mathcal{C}_t^\varepsilon)_{t\geqslant 0}$ on the state space $\{-1, 1\}$ and the

Q-matrix

$$Q^{\mathcal{C}}(t) = \begin{pmatrix} -\varphi_{\mathcal{C}}(t) & \varphi_{\mathcal{C}}(t) \\ \varphi_{\mathcal{C}}(t) & -\varphi_{\mathcal{C}}(t) \end{pmatrix}, \quad t \geqslant 0,$$

that for simplicity contains equal transition rates for both states given by

$$\varphi_{\mathcal{C}}(t) = p\left(\frac{t}{2T}\right) \varepsilon^{\alpha(t/2T)},$$

where $p\colon \mathbb{R}_+ \to [p_*, p^*]$ with $0 < p_* < p^* < \infty$ is piecewise-continuous and satisfies $p(t) = p(t+1)$ and α is given as in assumption (A2). Let T_n denote the n-th jump time of the Markov chain and define the normalized times $\tau_n = \frac{T_n}{T}$. The computation of the conditional density of τ_n given that $\tau_{n-1} = s \geqslant 0$ works similar as in Lemma 3.7 with φ_C replaced by $\varphi_{\mathcal{C}}$ and without a dependence on the value of $\mathcal{C}$ before the jump. The assertion about the Laplace transform can also partly be copied from Lemma 3.8. The functions L_- and L_+ coincide.

Lemma 4.1. *The conditional Laplace transform L of τ_n given $\tau_{n-1} = s \geqslant 0$ is*

$$L(x) = \frac{\int_s^{s+2} Tp\left(\frac{t}{2}\right) \varepsilon^{\alpha(t/2)} \,\mathrm{e}^{-xt - T\int_s^t p\left(\frac{r}{2}\right)\varepsilon^{\alpha(r/2)}\mathrm{d}r}\mathrm{d}t}{1 - \mathrm{e}^{-2x - T\int_0^2 p\left(\frac{r}{2}\right)\varepsilon^{\alpha(r/2)}\mathrm{d}r}}, \quad x \geqslant 0.$$

Assume $T = \varepsilon^{-\mu}$ for $\mu > 0$ and distinguish the following cases (see Figure 4.1).

(i) If $\mu \leqslant \alpha_*$ then the weak limit of the conditional law of τ_n is the null measure. This means $\lim_{\varepsilon\to 0} \mathbb{P}(\tau_n \geqslant t | \tau_{n-1} = s) = 1$ for all $t \geqslant s$.

(ii) If $\mu > \alpha^*$ then the conditional Laplace transform $L(x)$ converges to the Laplace transform of the Dirac measure in s.

(iii) Assume $\mu \in (\alpha_*, \alpha^*]$ and define $u \in [s, s+2)$ by $u = \inf\left\{t \geqslant s : \alpha\left(\frac{t}{2}\right) = \mu\right\}$. If additionally $\mu > \alpha\left(\frac{s}{2}\right)$ holds, the limit of the Laplace transform $L(x)$ is again the Dirac measure in s. For $\mu \leqslant \alpha\left(\frac{s}{2}\right)$ the weak limit of the conditional law of τ_n is the Dirac measure in u.

Proof. Mimic the first part of the proof of Lemma 3.8 to verify the formula of $L(x)$.
(i) In the case of $\mu < \alpha_*$ the integral $\int_0^1 p(r)\varepsilon^{\alpha(r)-\mu}\,\mathrm{d}r$ is bounded from above by $p^*\varepsilon^{\alpha_*-\mu}$ and thus converges to zero. That is why the numerator of $L(x)$ admitting this integral as an upper bound tends to zero and the denominator converges to $1 - \mathrm{e}^{-2x}$. All in all $L(x)$ is a null sequence as ε decreases. Let T be equal to $\varepsilon^{-\alpha_*}$ then from Laplace's method (Lemma 2.47) we can deduce

$$\lim_{\varepsilon\to 0} \sqrt{|\log\varepsilon|}\frac{\sqrt{\alpha''(a)}}{\sqrt{2\pi}p(a)} \int_0^1 p(t)\varepsilon^{\alpha(t)-\alpha_*}\mathrm{d}t = 1.$$

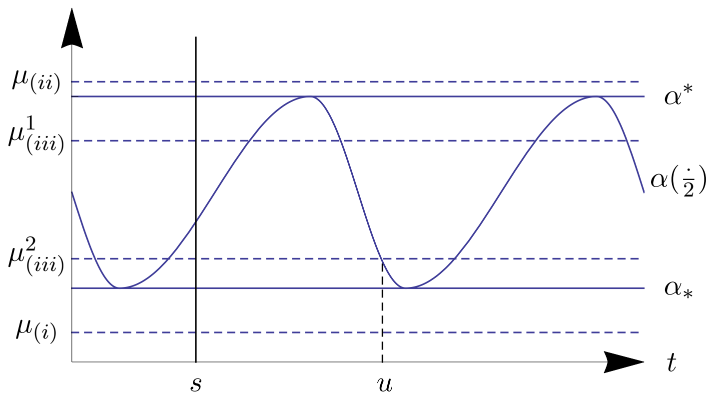


Figure 4.1: The values $\mu_{(i)}$, $\mu_{(ii)}$ and $\mu^1_{(iii)}$ and $\mu^2_{(iii)}$ correspond to the three cases considered in Lemma 4.1.

Due to Laplace's method (Lemma 2.47) for all $t \geqslant 0$ the integral

$$\int_t^{t+2} p\left(\frac{r}{2}\right) \varepsilon^{\alpha(r/2)-\alpha_*}\,\mathrm{d}r$$

is of order $|\log \varepsilon|^{-1/2}$ and thus tends to zero as ε does. The denominator of the Laplace transform $L(x)$ tends to $1-\mathrm{e}^{-2x}$ and the numerator also vanishes as ε converges to zero. This guarantees the null measure as weak limit for the conditional law of τ_n in the case of $T = \varepsilon^{-\alpha_*}$.

(ii) Assume $\mu > \alpha^*$. For all $t \geqslant 0$ we have

$$\lim_{\varepsilon\to 0} \int_t^{t+2} p\left(\frac{r}{2}\right) \varepsilon^{\alpha(r/2)-\mu}\,\mathrm{d}r \geqslant \lim_{\varepsilon\to 0} 2p_*\varepsilon^{\alpha^*-\mu} = \infty.$$

Thus the denominator of $L(x)$ tends to 1. Because of the last estimate for $t = s$ and

$$Tp\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)} \leqslant Tp\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)} + x, \quad x \geqslant 0$$

the numerator of the Laplace transform is bounded from above by e^{-xs} since we have

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\int_s^{s+2} p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\mu}\,\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r}\,\mathrm{d}t\\
&\leqslant \lim_{\varepsilon\to 0}\int_s^{s+2}\left(p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\mu}+x\right)\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r}\,\mathrm{d}t\\
&= \lim_{\varepsilon\to 0}\mathrm{e}^{-xs} - \mathrm{e}^{-x(s+2)-\int_s^{s+2} p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r}\\
&= \mathrm{e}^{-xs}.
\end{aligned}$$

It remains to verify the same lower bound. Estimate

$$
\begin{aligned}
&\lim_{\varepsilon\to 0}\int_s^{s+2} p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\mu}\,\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r}\,\mathrm{d}t\\
&\geqslant \lim_{\varepsilon\to 0}\int_s^{s+2}\frac{p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\mu}}{p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\mu}+x}\left(p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\mu}+x\right)\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r}\,\mathrm{d}t\\
&\geqslant \lim_{\varepsilon\to 0}\frac{1}{1+xp_*^{-1}\varepsilon^{\mu-\alpha^*}}\left(\mathrm{e}^{-xs}-\mathrm{e}^{-x(s+2)-\int_s^{s+2}p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r}\right)\\
&=\mathrm{e}^{-xs}.
\end{aligned}
$$

(iii) Assume $\mu\in(\alpha_*,\alpha^*]$. Due to $\mu>\alpha_*$ the exponent $\alpha\left(\frac{t}{2}\right)-\mu<0$ on an interval $B_t\subseteq[t,t+2]$, $t\geqslant 0$, of positive length. That ist why $\int_0^2 p\left(\frac{r}{2}\right)\varepsilon^{\alpha(r/2)-\mu}\,\mathrm{d}r$ diverges and we obtain the convergence to 1 of the denominator of $L(x)$. Concentrate on the special case $\mu=\alpha^*$ and assume without loss of generality $\alpha^*=\alpha\left(\frac{s}{2}\right)$. Calculate

$$
\begin{aligned}
&\int_s^{s+2} p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\alpha^*}\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\alpha^*}\mathrm{d}r}\,\mathrm{d}t\\
&=\int_s^{s+2}\left(-\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\alpha^*}\mathrm{d}r}\right)\mathrm{d}t-x\int_s^{s+2}\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\alpha^*}\mathrm{d}r}\,\mathrm{d}t\\
&=\left(\mathrm{e}^{-xs}-\mathrm{e}^{-x(s+2)-\int_s^{s+2}p(r/2)\varepsilon^{\alpha(r/2)-\alpha^*}\mathrm{d}r}\right)-x\int_s^{s+2}\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\alpha^*}\mathrm{d}r}\,\mathrm{d}t.
\end{aligned}
$$

The term in brackets tends to e^{-xs} because the integral of $p\left(\frac{r}{2}\right)\varepsilon^{\alpha(r/2)-\alpha^*}$ over $[s,s+2]$ diverges since $\alpha(r)-\alpha^*<0$ on a subset of $[s,s+2]$ of positive length. It remains to prove the convergence to zero of the integral of the exponential function. For any $\delta>0$ we estimate

$$
\begin{aligned}
\int_s^{s+2}\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\alpha^*}\mathrm{d}r}\mathrm{d}t &\leqslant 2\delta+\int_{s+\delta}^{s+2-\delta}\mathrm{e}^{-xt-(t-s)p_*\varepsilon^{\alpha((s+\delta)/2)\vee\alpha((s+2-\delta)/2)-\alpha^*}}\,\mathrm{d}t\\
&\leqslant 2\delta+\frac{2}{x+p_*\varepsilon^{\alpha((s+\delta)/2)\vee\alpha((s+2-\delta)/2)-\alpha^*}}
\end{aligned}
$$

which is small if $\varepsilon<\varepsilon_0(\delta)$, as long as $\alpha((s+\delta)/2)\vee\alpha((s+2-\delta)/2)<\alpha^*$. This finishes the proof of the convergence of $L(x)$ to e^{-xs} if $\mu=\alpha^*$. Now consider $\alpha^*>\mu>\alpha\left(\frac{s}{2}\right)\wedge\alpha_*$. Due to the inequality $\mu>\alpha_*$ the divergence of the integral of $p\left(\frac{r}{2}\right)\varepsilon^{\alpha(r/2)-\mu}$ over $[s,s+2]$ follows and yields the convergence to one of the denominator of $L(x)$. The numerator admits the upper bound e^{-xs}. Analogously as in (ii) we get for small $\delta>0$ that

$$
\begin{aligned}
&\lim_{\varepsilon\to 0}\int_s^{s+2} p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\mu}\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r}\mathrm{d}t\\
&\geqslant \lim_{\varepsilon\to 0}\int_s^{u-\delta}\frac{1}{1+xp_*^{-1}\varepsilon^{\mu-\alpha((u-\delta)/2)}}\left(-\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r}\right)\mathrm{d}t\\
&\geqslant \lim_{\varepsilon\to 0}\frac{1}{1+xp_*^{-1}\varepsilon^{\mu-\alpha((u-\delta)/2)}}\left(\mathrm{e}^{-xs}-\mathrm{e}^{-x(u-\delta)-\int_s^{u-\delta}p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r}\right)\\
&=\mathrm{e}^{-xs}
\end{aligned}
$$

due to $\alpha(\frac{t}{2})<\mu$ for $t\in[s,u-\delta]$. This proves the convergence to the Dirac measure in s.

Now consider the case when $\mu \in \left(\alpha_*, \alpha\left(\frac{s}{2}\right)\right]$. Again the denominator of $L(x)$ tends to 1. The analysis of the numerator demands a splitting of $[s, s+2]$ as on page 20 of [19]. Define $l_\varepsilon = \frac{\log\sqrt{|\log\varepsilon|}}{\log\varepsilon} < 0$ with $l_\varepsilon \to 0$ and

$$\Delta_1 = \left\{t \in [s, s+2] : \alpha\left(\frac{t}{2}\right) \geqslant \mu - l_\varepsilon\right\},$$
$$\Delta_2 = \left\{t \in [s, s+2] : \alpha\left(\frac{t}{2}\right) \leqslant \mu\right\},$$
$$\Delta_3 = \left\{t \in [s, s+2] : \alpha\left(\frac{t}{2}\right) \in [\mu, \mu - l_\varepsilon]\right\}.$$

The numerator of $L(x)$ equals to the sum of the integrals over Δ_1, Δ_2 and Δ_3. The integral over Δ_2 is decisive and the other two integrals are less important because

$$\lim_{\varepsilon\to 0}\int_{\Delta_1} p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\mu}\,\mathrm{e}^{-xt-\int_s^t \varepsilon^{\alpha(r/2)-\mu}p(r/2)\mathrm{d}r}\,\mathrm{d}t \leqslant \lim_{\varepsilon\to 0} 2p^*\varepsilon^{-l_\varepsilon} = \lim_{\varepsilon\to 0}\frac{2p^*}{\sqrt{|\log\varepsilon|}} = 0$$

and

$$\lim_{\varepsilon\to 0}\int_{\Delta_3} p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\mu}\,\mathrm{e}^{-xt-\int_s^t \varepsilon^{\alpha(r/2)-\mu}p(r/2)\mathrm{d}r}\,\mathrm{d}t \leqslant \lim_{\varepsilon\to 0} p^*|\Delta_3| = \lim_{\varepsilon\to 0} c\,|l_\varepsilon| = 0.$$

hold for some $c > 0$. Define $v \in (s, s+2]$ with $v = \inf\left\{t > u : \alpha\left(\frac{t}{2}\right) = \mu\right\}$ and calculate the integral over Δ_2 as below

$$\begin{aligned}
&\int_{\Delta_2} p\left(\frac{t}{2}\right)\varepsilon^{\alpha(t/2)-\mu}\,\mathrm{e}^{-xt-\int_s^t \varepsilon^{\alpha(r/2)-\mu}p(r/2)\mathrm{d}r}\,\mathrm{d}t\\
&= \int_u^v\left(-\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{-xt-\int_s^t \varepsilon^{\alpha(r/2)-\mu}p(r/2)\mathrm{d}r}\right)\mathrm{d}t - x\int_u^v \mathrm{e}^{-xt-\int_s^t \varepsilon^{\alpha(r/2)-\mu}p(r/2)\mathrm{d}r}\,\mathrm{d}t\\
&= \mathrm{e}^{-ux-\int_s^u p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r} - \mathrm{e}^{-vx-\int_s^v p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r} - x\int_u^v \mathrm{e}^{-xt-\int_s^t \varepsilon^{\alpha(r/2)-\mu}p(r/2)\mathrm{d}r}\,\mathrm{d}t.
\end{aligned}$$

Since $\alpha\left(\frac{r}{2}\right) - \mu < 0$ for $r \in (u, v)$ we have

$$\lim_{\varepsilon\to 0}\int_s^v p\left(\frac{r}{2}\right)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r = \infty.$$

Laplace's method (see Section 2.5) for the case where the minimal value of the function in the exponential is attained at an edgepoint of the considered interval guarantees that

$$\int_s^u p\left(\frac{r}{2}\right)\varepsilon^{\alpha(r/2)-\mu}\,\mathrm{d}r = \int_s^u p\left(\frac{r}{2}\right)\mathrm{e}^{-|\log\varepsilon|(\alpha(r/2)-\mu)}\,\mathrm{d}r = \mathrm{O}\left(|\log\varepsilon|^{-1}\right) \to 0$$

as ε converges to zero. Again it remains to verify the smallness of the integral over the exponential function as ε vanishes. Estimate

$$\begin{aligned}
\int_u^v \mathrm{e}^{-xt-\int_s^t p(r/2)\varepsilon^{\alpha(r/2)-\mu}\mathrm{d}r}\,\mathrm{d}t &\leqslant 2\delta + \int_{u+\delta}^{v-\delta} \mathrm{e}^{-xt-(t-s)p_*\varepsilon^{\alpha((u+\delta)/2)\vee\alpha((v-\delta)/2)-\mu}}\,\mathrm{d}t\\
&\leqslant 2\delta + \frac{2}{x + p_*\varepsilon^{\alpha((u+\delta)/2)\vee\alpha((v-\delta)/2)-\mu}}
\end{aligned}$$

which is small if ε is chosen sufficiently small because $\alpha((u+\delta)/2) \vee \alpha((v-\delta)/2) - \mu < 0$. Overall the denominator of $L(x)$ tends to e^{-xu} if $\mu \in (\alpha_*, \alpha\left(\frac{s}{2}\right)]$. $\square$

Time scales $T = \varepsilon^{-\mu}$ for $\mu \geqslant \alpha^*$ are much too big to distinguish jumps. If $\mu \leqslant \alpha_*$ we do not see any jump with high probability and the particle is trapped in the initial state. For $\mu \in (\alpha_*, \alpha^*)$ the Markov chain changes between a regime of chaos where the stability index is smaller than μ and a regime of trapping if $\alpha(t)$ is bigger than μ. The correct time scale to separate jumps must lie somewhere between $\varepsilon^{-\alpha_*}$ and $\varepsilon^{-\alpha^*}$. Since the jump affinity of $\mathcal{C}$ is the greatest for such t with $\alpha(t)$ minimal, it would be natural to think $\varepsilon^{-\alpha_*}$ is the optimal rate. But then there is not enough time for the process to perform the desired jumps. A rate slightly bigger than $\varepsilon^{-\alpha_*}$ would overcome this problem. The probabilistic measure of optimal tuning already used in Section 3.1 helps to reveal the correct time scale.

Proposition 4.2. *Fix $\delta > 0$ small and assume $k \in \mathbb{N}_0$. Then $T \mapsto \mathbb{P}_{-1}(\tau_1 \in [2(k+a)-\delta, 2(k+a)+\delta])$ attains its largest value at the optimal half-period length*

$$T_{k,\delta}(\varepsilon) = \frac{1}{\int_{2(k+a)-\delta}^{2(k+a)+\delta} p\left(\frac{r}{2}\right) \varepsilon^{\alpha(r/2)}\, \mathrm{d}r} \log \frac{\int_0^{2(k+a)+\delta} p\left(\frac{r}{2}\right) \varepsilon^{\alpha(r/2)}\, \mathrm{d}r}{\int_0^{2(k+a)-\delta} p\left(\frac{r}{2}\right) \varepsilon^{\alpha(r/2)}\, \mathrm{d}r}.$$

As $\delta \to 0$ the optimal half-period length $T_{k,\delta}(\varepsilon)$ tends to $T_k(\varepsilon) = \left(\int_0^{2(k+a)} p\left(\frac{r}{2}\right) \varepsilon^{\alpha(r/2)}\, \mathrm{d}r\right)^{-1}$.

Proof. Due to the similarity of the result to Propositions 3.5 and 3.9 we omit the proof. □

Corollary 4.3. *Under the assumptions of the previous proposition we have*

$$\lim_{\varepsilon \to 0} \frac{T_k(\varepsilon)\varepsilon^{\alpha_*}}{\sqrt{|\log \varepsilon|}} \frac{\sqrt{2\pi}\, p(a)(2k+1)}{\sqrt{\alpha''(a)}} = 1.$$

Proof. Apply the periodicity of α and p to obtain

$$\int_0^{2k+2a} p\left(\frac{r}{2}\right) \varepsilon^{\alpha(r/2)}\, \mathrm{d}r = 2k \int_0^1 p(r)\varepsilon^{\alpha(r)}\, \mathrm{d}r + 2\int_0^a p(r)\varepsilon^{\alpha(r)}\, \mathrm{d}r.$$

Laplace's method (Lemma 2.47) implies that the first integral is of order $\frac{\sqrt{2\pi}p(a)}{\sqrt{\alpha''(a)}}\varepsilon^{\alpha_*}|\log \varepsilon|^{-1/2}$ and the last one is of order $\frac{1}{2}\frac{\sqrt{2\pi}p(a)}{\sqrt{\alpha''(a)}}\varepsilon^{\alpha_*}|\log \varepsilon|^{-1/2}$ as ε tends to zero. □

4.2 The periodic jump diffusion and the decomposition of jumps

The previous section serves as simplification of the dynamics of the solution Y^ε of (4.1) which is now at the centre of attention.

In comparison to the previous chapter the basins of attraction are time-independent here.

Definition 4.4. *Assume $y_t(y)$ solves $\frac{\mathrm{d}}{\mathrm{d}t} y_t = -\nabla V(y_t)$ with initial value $y \in \mathbb{R}^d$. Define*

$$\Omega_\pm = \left\{ y \in \mathbb{R}^d : \lim_{t\to\infty} y_t(y) = m_\pm \right\}.$$

Let $\Gamma = \mathbb{R}^d \backslash (\Omega_+ \cup \Omega_-)$ denote the separatrix between Ω_- and Ω_+.

Definition 4.5. *Assume $\delta_1 > \delta_2 > 0$ and define the reduced domains of attraction by*

$$\Omega_\pm(\delta_1) = \{y \in \Omega_\pm : B_{\delta_1}(y_t(y)) \subseteq \Omega_\pm, \text{ for all } t \geqslant 0\},$$
$$\Omega_\pm(\delta_1,\delta_2) = \{y \in \Omega_\pm(\delta_1) : B_{\delta_2}(y_t(y)) \subseteq \Omega_\pm(\delta_1), \text{ for all } t \geqslant 0\},$$
$$\Omega_\pm(\delta_1,\delta_2,\delta_2) = \{y \in \Omega_\pm(\delta_1,\delta_2) : B_{\delta_2}(y_t(y)) \subseteq \Omega_\pm(\delta_1,\delta_2), \text{ for all } t \geqslant 0\},$$

and the enlarged separatrices by

$$\Gamma(\delta_1) = \mathbb{R}^d \setminus (\Omega_+(\delta_1) \cup \Omega_-(\delta_1)),$$
$$\Gamma(\delta_1,\delta_2) = \mathbb{R}^d \setminus (\Omega_+(\delta_1,\delta_2) \cup \Omega_-(\delta_1,\delta_2)),$$
$$\Gamma(\delta_1,\delta_2,\delta_2) = \mathbb{R}^d \setminus (\Omega_+(\delta_1,\delta_2,\delta_2) \cup \Omega_-(\delta_1,\delta_2,\delta_2)).$$

Recall the union $O_L^V = \bigcup_{c\leqslant L}\{y \in \mathbb{R}^d : V(y) = c\}$ of all level sets of V below L for $L \geqslant R^* := R_V^*$ given in (V3) *and define $O_L = O_L^V$ which is an invariant set for the dynamical system $\dot{y} = -\nabla V(y)$. Intersection of O_L with all sets given above defines the bounded and reduced domains of attraction $\Omega_{\pm,L}(\delta_1)$, $\Omega_{\pm,L}(\delta_1,\delta_2)$, and $\Omega_{\pm,L}(\delta_1,\delta_2,\delta_2)$ and the truncated and enlarged separatrices $\Gamma_L(\delta_1)$, $\Gamma_L(\delta_1,\delta_2)$ and $\Gamma_L(\delta_1,\delta_2,\delta_2)$. Let $\Omega_{\pm,L}$ be the bounded set $\Omega_\pm \cap O_L$ and Γ_L denotes the truncated separatrix $\Gamma \cap O_L$.*

Lemma 4.6. *For all $\delta > 0$, $L \geqslant R^*$ and $\gamma > 0$ there are $C_{log}^\pm = C_{log}^\pm(\gamma,\delta,L) > 0$ and $\varepsilon_0^\pm = \varepsilon_0^\pm(\gamma,\delta,L) > 0$ such that*

$$\|y_t(y_0) - m_\pm\| \leqslant \frac{1}{2}\varepsilon^{2\gamma}$$

is true for all $t \geqslant \gamma C_{log}^\pm|\log\varepsilon|$ and $y_0 \in \Omega_{\pm,L}(\delta)$. Define $R_{\varepsilon^\gamma} := (C_{log}^+ \vee C_{log}^-)\gamma|\log\varepsilon|$.

Proof. Due to the steepness of V and the Lemma of Gronwall the solution y starting in $\Omega_{\pm,L}(\delta)$ reaches a small neighbourhood of $m_\pm$ in a finite time depending on L, δ and the choice of the well. In the vicinity of $m_\pm$ we apply the Lyapunov method to calculate the current rate of convergence. Define $\bar{y}_t = y_t - m_\pm$ which satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\bar{y}_t = -\nabla V(\bar{y}_t + m_\pm).$$

A linearization of the system allows to approximate the asymptotic behaviour of $\bar{y}_t$ by the behaviour of the solution of the linear differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\bar{\bar{y}}_t = \left(-\frac{\partial^2}{\partial\bar{\bar{y}}_i\partial\bar{\bar{y}}_j}V(m_\pm)\right)_{i,j=1}^d \bar{\bar{y}}_t.$$

Since the eigenvalues of the Hesse matrix of V at $m_\pm$ are all positive the solution $\bar{y} \equiv 0$ is uniformly asymptotically stable (Theorem 15.3 in [1]). Finally y_t belongs to $B_\varepsilon(m_\pm)$ for all $t \geqslant c|\log\varepsilon|$ for some $c > 0$. □

Decomposition of jumps:

As in the previous chapter our methods demand an ε-dependent decomposition of jumps.

Definition 4.7. *Let c_A in the decomposition (2.4) of A^T be equal to $\varepsilon^{-\rho}$ for some arbitrary $\rho \in (0,1)$. The sequence $(\tau_j)_{j\in\mathbb{N}_0}$ with $\tau_0 = 0$ denotes the jump times of the big jump part $A^T - \widetilde{A}^T$ and $(W_j)_{j\in\mathbb{N}}$ are the corresponding jump sizes, while εW_j admits values in $B^c_{\varepsilon^{1-\rho}}(0)$. Define $\beta^\varepsilon(t) = \nu_t(B^c_{\varepsilon^{-\rho}}(0))$.*

From assumption (A2) and the substitution $y = \varepsilon^\rho x$ we immediately derive

$$\beta^\varepsilon(t) = \nu_t(B^c_{\varepsilon^{-\rho}}(0)) = \int_{\|x\|\geqslant\varepsilon^{-\rho}} \frac{c(t)}{\|x\|^{d+\alpha(t)}}\,\mathrm{d}x = \int_{\|y\|\geqslant 1} \frac{c(t)\varepsilon^{\rho\alpha(t)}}{\|y\|^{d+\alpha(t)}}\,\mathrm{d}y = \varepsilon^{\rho\alpha(t)}\nu_t(B^c_1(0)).$$

The conditional law of εW_j given $\tau_j = t$ equals to $\nu_{t/2T}\left(B^c_{\varepsilon^{-\rho}}(0) \cap \frac{1}{\varepsilon}\cdot\right)/\nu_{t/2T}(B^c_{\varepsilon^{-\rho}}(0))$.

Lemma 4.8. *The conditional density of the inter-jump time $\tau_{j+1} - \tau_j$ given $\tau_j = r \geqslant 0$ is*

$$f_{\tau_{j+1}-\tau_j|\tau_j=r}(t) = \beta^\varepsilon\left(\frac{r+t}{2T}\right)\exp\left(-\int_r^{r+t}\beta^\varepsilon\left(\frac{s}{2T}\right)\mathrm{d}s\right),$$

for $t \geqslant 0$ and $j \in \mathbb{N}_0$. In particular one obtains

$$f_{\tau_1}(t) = \beta^\varepsilon\left(\frac{t}{2T}\right)\exp\left(-\int_0^t \beta^\varepsilon\left(\frac{s}{2T}\right)\mathrm{d}s\right), \quad t \geqslant 0.$$

Proof. From assumption (A2) we can deduce

$$\begin{aligned}
\mathbb{P}(\tau_{j+1} - \tau_j \geqslant t|\tau_j = r) &= \mathbb{P}\left(N^{A^T}(B^c_{\varepsilon^{-\rho}}(0) \times (r, r+t]) = 0\right) \\
&= \exp\left(-\int_r^{r+t}\nu_{s/2T}(B^c_{\varepsilon^{-\rho}}(0))\,\mathrm{d}s\right) \\
&= \exp\left(-\int_r^{r+t}\beta^\varepsilon\left(\frac{s}{2T}\right)\mathrm{d}s\right).
\end{aligned}$$

We can calculate the distributional function and immediately get the density

$$f_{\tau_{j+1}-\tau_j|\tau_j=r}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(1 - \exp\left(-\int_r^{r+t}\beta^\varepsilon\left(\frac{s}{2T}\right)\mathrm{d}s\right)\right).$$

From differentiation we can deduce the formula given in the lemma. □

4.3 The small jump process

To a great extent the methods used for the small jump process in the previous chapter can be transferred to the current setting. Sometimes just little modifications concerning special features of additive processes have to be made.

Definition 4.9. *Let $s \geqslant 0$. Assume $(\widetilde{Y}^\varepsilon_{s,t}(y))_{t\geqslant 0}$ denotes the solution of*

$$\widetilde{Y}^\varepsilon_{s,t}(y) = y - \int_0^t \nabla V(\widetilde{Y}^\varepsilon_{s,r})\,\mathrm{d}r + \varepsilon\left(\widetilde{A}^T_{s+t} - \widetilde{A}^T_s\right), \quad t \geqslant 0.$$

Let $y_t(y)$ be the solution of $\frac{\mathrm{d}}{\mathrm{d}t}y_t = -\nabla V(y_t)$ with $y_0 = y$.

We suggest that on an interval of the length of an inter-jump period of the big jump part of A^T, the small jump process hardly deviates from the deterministic solution. The proof is prepared in the remaing part of this section.

Theorem 4.10. *Assume $L \geqslant R^*$ and $\delta_0 > 0$. Let T_r, for $r > 0$, denote a random variable with density*

$$f_r(t) = \beta^\varepsilon\left(\frac{r+t}{2T}\right)\exp\left(-\int_r^{r+t}\beta^\varepsilon\left(\frac{s}{2T}\right)\mathrm{d}s\right), \quad t \geqslant 0.$$

Then there are $\varepsilon_0, p_0, \gamma_0 > 0$ such that the following inequality is valid for all $\varepsilon < \varepsilon_0, p < p_0$ and $\gamma < \gamma_0$

$$\sup_{r\geqslant 0}\sup_{y\in\Omega_{\pm,L}(\delta_0)}\mathbb{P}\left(\sup_{t\in[0,T_r]}\left\|\widetilde{Y}^\varepsilon_{r,t}(y) - y_t(y)\right\| \geqslant \varepsilon^\gamma\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}.$$

4.3.1 The martingale estimate and the boundedness of the solution

The local martingale part $\widetilde{A}^M$ of $\widetilde{A}^T$ that is square-integrable (cf. Lemma 2.19) is given by

$$\widetilde{A}^M_t = \int_0^t \sigma(s)\,\mathrm{d}W_s + \int_0^t\int_{\|x\|<\varepsilon^{-\rho}} x\,\widetilde{N}^{A^T}(\mathrm{d}x, \mathrm{d}s). \tag{4.2}$$

It is still depending on T but we omit T for simplicity of the notation.

Lemma 4.11. *Let $(X_t)_{t\geqslant 0}$ be a d-dimensional, $(\mathcal{F}_{s+t})_{t\geqslant 0}$-adapted, bounded, and càdlàg process. Then there exist $\varepsilon_0, p_0, \beta_0, \theta_0 > 0$ such that the following inequality*

$$\sup_{s\geqslant 0}\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}}\varepsilon\Big|\sum_{j=1}^d\int_s^{s+t}X^j_{r-}\,\mathrm{d}\widetilde{A}^{M,j}_r\Big| \geqslant \varepsilon^\beta\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

holds for all $\varepsilon \in (0, \varepsilon_0), \beta \in (0, \beta_0), \theta \in (0, \theta_0)$ and $p \in (0, p_0)$.

Proof. As in the proof of Lemma 3.20 we deal with a local martingale with bounded jumps given by

$$\mathcal{M}^s_t = \sum_{j=1}^d\int_s^{s+t}X^j_{r-}\,\mathrm{d}\widetilde{A}^{M,j}_r.$$

By repeating the method of Kallenberg (proof of Theorem 23.17 in [30]) as used in the proof of Lemma 3.20 it remains to verify

$$\sup_{s\geqslant 0}\mathbb{P}\left(\varepsilon^2\left[\mathcal{M}^s\right]_{\varepsilon^{-\theta}} \geqslant \varepsilon^{4\beta}\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}.$$

From the orthogonality of the continuous and the pure jump part of $\mathcal{M}^s$ we derive the succeeding decomposition for the quadratic variation of $\mathcal{M}^s$

$$\varepsilon^2\left[\mathcal{M}^s\right]_t = \varepsilon^2\Big[\sum_{j=1}^d\int_s^{s+\cdot}\sum_{i=1}^d X^i_r\sigma_{ij}(r)\,\mathrm{d}W^j_r\Big]_t + \Big[\int_s^{s+\cdot}\int_{\|x\|\in(0,\varepsilon^{-\rho})}\varepsilon\sum_{j=1}^d X^j_{r-}x_j\,\widetilde{N}^{A^T}(\mathrm{d}x, \mathrm{d}r)\Big]_t$$

for $t \geqslant 0$. The quadratic variation of the Gaussian part is calculated explicitly as

$$\Big[\sum_{j=1}^{d} \int_s^{s+\cdot} \sum_{i=1}^{d} X_r^i \sigma_{ij}(r)\,\mathrm{d}W_r^j\Big]_t = \sum_{j,k=1}^{d} \int_s^{s+t} \left(\sum_{i=1}^{d} X_r^i \sigma_{ij}(r)\right)\left(\sum_{i=1}^{d} X_r^i \sigma_{ik}(r)\right) \mathrm{d}[W^j, W^k]_r$$
$$= \int_s^{s+t} \sum_{j=1}^{d} \left(\sum_{i=1}^{d} X_r^i \sigma_{ij}(r)\right)^2 \mathrm{d}r,$$

for $t \geqslant 0$. Thus from the boundedness of X and all components of σ and Lemma 2.21 we can deduce that

$$\varepsilon^2 \left[\mathcal{M}^s\right]_{\varepsilon^{-\theta}} \leqslant c_1 \varepsilon^{2-\theta} + \int_s^{s+\varepsilon^{-\theta}} \int_{\|x\| \in (0,\varepsilon^{-\rho})} \varepsilon^2 \Big(\sum_{j=1}^{d} X_{r-}^j x_j\Big)^2 N^{A^T}(\mathrm{d}x, \mathrm{d}r)$$
$$\leqslant c_1 \varepsilon^{2-\theta} + \varepsilon^2 c_2 \sum_{t \leqslant \varepsilon^{-\theta}} \|\Delta \widetilde{A}_{s+t}^T\|^2,$$

for some $c_1, c_2 > 0$. Proving the next lemma which is the counterpart of Lemma 3.19 for additive processes finishes this proof. □

Lemma 4.12. *There are $\varepsilon_0, \theta_0, \beta_0 > 0$ and $p_0 > 0$ that guarantee the inequality*

$$\sup_{r \geqslant 0} \mathbb{P}\left(\varepsilon^2 \sum_{t \leqslant \varepsilon^{-\theta}} \|\Delta \widetilde{A}_{r+t}^T\|^2 \geqslant \varepsilon^\beta\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

for all $\varepsilon \in (0, \varepsilon_0), \beta \in (0, \beta_0), \theta \in (0, \theta_0)$ and $p \in (0, p_0)$.

Proof. As in the proof of Lemma 3.19 the Markov inequality transfers the problem to the verification of the finiteness of the expectation

$$\mathbb{E} \exp\left(\varepsilon^{2-2\beta} \sum_{t \leqslant \varepsilon^{-\theta}} \|\Delta \widetilde{A}_{r+t}^T\|^2\right).$$

We suggest the equality

$$\mathbb{E}\, \mathrm{e}^{-\lambda \sum_{s \leqslant t} \|\Delta \widetilde{A}_{r+s}^T\|^2} = \exp\left(-\int_r^{r+t} \int_{\|x\| \leqslant \varepsilon^{-\rho}} \left(1 - \mathrm{e}^{-\lambda \|x\|^2}\right) \nu_{s/2T}(\mathrm{d}x)\, \mathrm{d}s\right) \tag{4.3}$$

for all $r, t \geqslant 0$ and $\lambda \in \mathbb{R}$. With $\lambda = -\varepsilon^{2-2\beta}$ and $t = \varepsilon^{-\theta}$ this assertion allows us to estimate as in the stated proof

$$\mathbb{E} \exp\left(\varepsilon^{2-2\beta} \sum_{t \leqslant \varepsilon^{-\theta}} \|\Delta \widetilde{A}_{r+t}^T\|^2\right)$$
$$\leqslant \exp\left(2\varepsilon^{2-2\beta-\theta} \max_{t \in [0,1]} \int_{\|x\| \in (0,1)} \|x\|^2\, \nu_t(\mathrm{d}x) + \left(\mathrm{e}^{\varepsilon^{2-2\beta-2\rho}} - 1\right) \varepsilon^{-\theta} \max_{t \in [0,1]} \nu_t(\|x\| \geqslant 1)\right),$$

where the latter exponent is bounded if $2 - 2\beta - 2\rho - \theta > 0$.

It remains to prove equation (4.3). An analogous result for a Lévy process falls back on the property of the sum of the squared jumps to be a Lévy subordinator but here the same methods as in Proposition 4.26 in Chapter 2 of [27] are helpful. For $r \geqslant 0$ and $\lambda > 0$ define

$$X_t^r = \int_r^{r+t} \int_{\|x\| \in (0, \varepsilon^{-\rho})} \lambda \|x\|^2 \, N^{A^T}(\mathrm{d}x, \mathrm{d}s),$$

and $E_t^r = \mathrm{e}^{-X_t^r}$. The Itô formula implies

$$E_t^r = 1 - \int_0^t E_{s-}^r \, \mathrm{d}X_s^r + \sum_{s \leqslant t} \left(E_s^r - E_{s-}^r + E_{s-}^r \Delta X_s^r \right) = 1 - \sum_{s \leqslant t} E_{s-}^r (1 - \mathrm{e}^{-\Delta X_s^r}).$$

Introducing the necessary compensator yields

$$\begin{aligned} E_t^r = 1 &- \int_r^{r+t} \int_{\|x\| \in (0, \varepsilon^{-\rho})} E_{(s-r)-}^r \left(1 - \mathrm{e}^{-\lambda \|x\|^2} \right) \widetilde{N}^{A^T}(\mathrm{d}x, \mathrm{d}s) \\ &- \int_r^{r+t} \int_{\|x\| \in (0, \varepsilon^{-\rho})} E_{(s-r)-}^r \left(1 - \mathrm{e}^{-\lambda \|x\|^2} \right) \nu_{s/2T}(\mathrm{d}x) \, \mathrm{d}s. \end{aligned}$$

For $\lambda > 0$ we have $|E_t^r| \leqslant 1$. Thus the integral with respect to $\widetilde{N}^{A^T}$ is a martingale because

$$\begin{aligned} &\mathbb{E} \left| \int_r^{r+t} \int_{\|x\| \in (0, \varepsilon^{-\rho})} E_{(s-r)-}^r \left(1 - \mathrm{e}^{-\lambda \|x\|^2} \right) \widetilde{N}^{A^T}(\mathrm{d}x, \mathrm{d}s) \right|^2 \\ &\leqslant \int_r^{r+t} \int_{\|x\| \in (0, \varepsilon^{-\rho})} \left(1 - \mathrm{e}^{-\lambda \|x\|^2} \right)^2 \nu_{s/2T}(\mathrm{d}x) \, \mathrm{d}s \end{aligned}$$

is finite. Through this result, the equality derived from the Itô formula above, and Fubini's Theorem we can compute the expectation of E_t^r as follows

$$\mathbb{E} E_t^r = 1 - \int_r^{r+t} \mathbb{E} E_{(s-r)-}^r \int_{\|x\| \in (0, \varepsilon^{-\rho})} \left(1 - \mathrm{e}^{-\lambda \|x\|^2} \right) \nu_{s/2T}(\mathrm{d}x) \, \mathrm{d}s.$$

With $f^r(t) = \mathbb{E} E_t^r$ this yields the simple differential equation

$$\mathrm{d}f^r(t) = -f^r(t) \int_{\|x\| \in (0, \varepsilon^{-\rho})} \left(1 - \mathrm{e}^{-\lambda \|x\|^2} \right) \nu_{(r+t)/2T}(\mathrm{d}x) \, \mathrm{d}t,$$

which admits the solution

$$\begin{aligned} f^r(t) &= \exp \left(- \int_0^t \int_{\|x\| \in (0, \varepsilon^{-\rho})} \left(1 - \mathrm{e}^{-\lambda \|x\|^2} \right) \nu_{(r+s)/2T}(\mathrm{d}x) \, \mathrm{d}s \right) \\ &= \exp \left(- \int_r^{r+t} \int_{\|x\| \in (0, \varepsilon^{-\rho})} \left(1 - \mathrm{e}^{-\lambda \|x\|^2} \right) \nu_{s/2T}(\mathrm{d}x) \, \mathrm{d}s \right). \end{aligned}$$

Formula (4.3) is proven for $\lambda > 0$. Due to the finiteness of the right-hand side of (4.3) for all $\lambda \in \mathbb{R}$ the definition of the Laplace transform is well-defined also for $\lambda \leqslant 0$. □

Analogously to Lemma 3.21 in Chapter 3 we need a boundedness result for the small jump process.

Lemma 4.13. *For all $L \geqslant R^*$ there exist $N \in \mathbb{N}$ and $\varepsilon_0, \theta_0, p_0 > 0$ such that*

$$\sup_{s\geqslant 0} \sup_{y_0 \in O_L} \mathbb{P}\left(\sup_{t\leqslant \varepsilon^{-\theta}} \left\|\widetilde{Y}^{\varepsilon}_{s,t}(y_0)\right\| \geqslant N \right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

is true for all $\varepsilon \in (0,\varepsilon_0), \theta \in (0,\theta_0)$ and $p \in (0,p_0)$.

Proof. Choose $s \geqslant 0$ and $y_0 \in O_L$ for $L \geqslant R^*$. Assume $\|\nabla \log(1+V(y))\| \leqslant C$ for all $y \in \mathbb{R}^d$ which is possible because of (V4). As in step three of the proof of Proposition 2.39 we have

$$V(y) \geqslant \frac{c_V^*}{C}\|y\|^{1+c_V^*} - 1$$

for all $y \in O^c_{R^*}$ due to assumption (V3). This guarantees $\log(1+V(y)) \geqslant u_n$ for $u_n := \log\left(\frac{c_V^*}{C} n^{1+c_V^*}\right)$ for $y \in O_n^c$ with $n \geqslant \lceil R^* \rceil$. The sequence $(u_n)_{n\in\mathbb{N}}$ diverges as $n \to \infty$. Choose $N \in \mathbb{N}$ such that $\log(1+V(y_0)) < \frac{u_N}{2}$ for all $y_0 \in O_L$ and define $f(y) = \log(1+V(y))$. The inclusion

$$\left\{ \sup_{t\leqslant\varepsilon^{-\theta}} \left\|\widetilde{Y}^{\varepsilon}_{s,t}(y_0)\right\| \geqslant N \right\} \subseteq \left\{ \sup_{t\leqslant\varepsilon^{-\theta}} \log\left(1+V\left(\widetilde{Y}^{\varepsilon}_{s,t}(y_0)\right)\right) \geqslant u_N \right\}$$

and Itô's formula prove

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} f(\widetilde{Y}^{\varepsilon}_{s,t}(y_0)) \geqslant u_N \right)\\
&\leqslant \mathbb{P}\left(f(y_0) \geqslant \frac{u_N}{2}\right) + \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon\left| \int_0^t \nabla f(\widetilde{Y}^{\varepsilon}_{s,r}) \cdot \gamma(s+r)\,\mathrm{d}r \right| \geqslant \varepsilon^{\beta} \right)\\
&\quad + \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon\left| \sum_{i=1}^d \int_s^{s+t} \frac{\partial}{\partial y_i} f(\widetilde{Y}^{\varepsilon}_{s,(r-s)-})\,\mathrm{d}\widetilde{A}^{M,i}_r \right| \geqslant \varepsilon^{\beta} \right)\\
&\quad + \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \frac{\varepsilon^2}{2}\left| \sum_{i,j=1}^d \int_0^t \frac{\partial^2}{\partial y_i \partial y_j} f(\widetilde{Y}^{\varepsilon}_{s,r})\,\mathrm{d}\left[\widetilde{Y}^{\varepsilon,i}_{s,\cdot}, \widetilde{Y}^{\varepsilon,j}_{s,\cdot}\right]^c_r \right| \geqslant \varepsilon^{\beta} \right)\\
&\quad + \mathbb{P}\left(\sup_{t\leqslant\varepsilon^{-\theta}} \varepsilon^2\left| \sum_{r\leqslant t} \left(f(\widetilde{Y}^{\varepsilon}_{s,r}) - f(\widetilde{Y}^{\varepsilon}_{s,r-}) - \nabla f(\widetilde{Y}^{\varepsilon}_{s,r-})\Delta\widetilde{Y}^{\varepsilon}_{s,r} \right) \right| \geqslant \varepsilon^{\beta} \right)\\
&=: I_1 + \ldots + I_5.
\end{aligned}$$

While $I_1 = 0$ due to the choice of N, the term I_2 is small since ∇f and γ are bounded (assumptions (V4) and (A1)). The third summand can be treated through Lemma 4.11. The summand I_4 requires knowledge about the continuous part of the quadratic variation of $\widetilde{Y}^{\varepsilon}$. It is given by

$$\left[\widetilde{Y}^{\varepsilon,i}_{s,\cdot}, \widetilde{Y}^{\varepsilon,j}_{s,\cdot}\right]^c_u = \sum_{k,l=1}^d \int_s^{s+u} \sigma_{ik}(r)\sigma_{jl}(r)\,\mathrm{d}[W^k, W^l]_r = \int_s^{s+u} \sum_{k=1}^d \sigma_{ik}(r)\sigma_{jk}(r)\,\mathrm{d}r, \quad u \geqslant 0$$

because W^k and W^l are independent for $k \neq l$. If $2-\theta-\beta > 0$ and ε is small, we obtain

$$\mathbb{P}\left(\sup_{t\in[0,\varepsilon^{-\theta}]} \left| \frac{\varepsilon^2}{2} \sum_{i,j,k=1}^d \int_0^t \frac{\partial^2}{\partial y_i \partial y_j} f(\widetilde{Y}^{\varepsilon}_{s,r})\sigma_{ik}(s+r)\sigma_{jk}(s+r)\,\mathrm{d}r \right| \geqslant \varepsilon^{\beta} \right) = 0.$$

On account of the boundedness of the derivatives of f of second order we get

$$\sum_{r\leqslant t}\Big(f(\widetilde{Y}^{\varepsilon}_{s,r})-f(\widetilde{Y}^{\varepsilon}_{s,r-})-\nabla f(\widetilde{Y}^{\varepsilon}_{s,r-})\Delta\widetilde{Y}^{\varepsilon,j}_{s,r}\Big)\leqslant c_1\varepsilon^2\sum_{r\leqslant t}\|\Delta\widetilde{A}^T_{s+r}\|^2$$

for some $c_1>0$. With Lemma 4.12 in mind the proof is complete. □

4.3.2 Behaviour until the logarithmic return time and near the well minima

Lemma 4.14. *For $L\geqslant R^*$ and $\delta_0>0$ there exist $\varepsilon_0,\beta_0>0$ and $p_0>0$ such that*

$$\sup_{r\geqslant 0}\sup_{y\in\Omega_{\pm,L}(\delta_0)}\mathbb{P}\left(\sup_{t\in[0,R_{\varepsilon^\beta}]}\Big\|\widetilde{Y}^{\varepsilon}_{r,t}(y)-y_t(y)\Big\|\geqslant\varepsilon^\beta\right)\leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

holds for all $\varepsilon<\varepsilon_0,\beta<\beta_0$ and $p<p_0$.

Proof. For $L\geqslant R^*$ choose N according to Lemma 4.13. On O_L the gradient of V is locally Lipschitz continuous with constant $C_L>0$. On the set $\Big\{\sup_{t\leqslant\varepsilon^{-\theta}}\|\widetilde{Y}^{\varepsilon}_{r,t}(y)\|<N\Big\}$ the small jump process is bounded and through the use of the Lemma of Gronwall we obtain

$$\|\widetilde{Y}^{\varepsilon}_{r,t}(y)-y_t(y)\|\leqslant \mathrm{e}^{C_Lt}\sup_{s\leqslant t}\varepsilon\|\widetilde{A}^T_{r+s}-\widetilde{A}^T_r\|.$$

Due to the inequality above, $\varepsilon^\beta\mathrm{e}^{-C_LR_{\varepsilon^\beta}}=\varepsilon^{\beta c_1}$ for some $c_1>0$ it remains to prove the existence of $\varepsilon_0,\zeta_0,p_0>0$ and $\theta_0>0$ such that the estimate

$$\sup_{r\geqslant 0}\mathbb{P}\left(\sup_{t\in[0,\varepsilon^{-\theta}]}\varepsilon\|\widetilde{A}^T_{t+r}-\widetilde{A}^T_r\|\geqslant\varepsilon^\zeta\right)\leqslant\mathrm{e}^{-\varepsilon^{-p}}$$

is valid for $\varepsilon\in(0,\varepsilon_0),\zeta\in(0,\zeta_0),p\in(0,p_0)$ and $\theta\in(0,\theta_0)$ because $R_{\varepsilon^\beta}<\varepsilon^{-\theta}$ for small ε. Then choose $\beta_0=\zeta_0c_1^{-1}$. The decomposition of $\widetilde{A}^T$ into a deterministic part and a local martingale part given in (4.2) implies

$$\|\widetilde{A}^T_{t+r}-\widetilde{A}^T_r\|\leqslant\varepsilon\Big\|\int_r^{r+t}\gamma(s)\,\mathrm{d}s\Big\|+\varepsilon\|\widetilde{A}^M_{r+t}-\widetilde{A}^M_r\|.$$

Due to the boundedness of γ given in assumption (A1) the first summand of the upper bound above is smaller than $\varepsilon^{1-\theta}\|\gamma\|\leqslant\varepsilon^{2\zeta}$ if $1-\theta-2\zeta>0$ and ε is small. The equivalence of norms in $\mathbb{R}^d$ and Lemma 4.11 verify the claimed result. □

Lemma 4.15. *There are $\varepsilon_0,\beta_0,\theta_0>0$ and $p_0>0$ such that*

$$\sup_{r\geqslant 0}\sup_{y:\|y-m_\pm\|\leqslant\varepsilon^{2\beta}}\mathbb{P}\left(\sup_{t\in[0,\varepsilon^{-\theta}]}\Big\|\widetilde{Y}^{\varepsilon}_{r,t}(y)-m_\pm\Big\|\geqslant\varepsilon^\beta\right)\leqslant\mathrm{e}^{-\varepsilon^{-p}},$$

is true for all $\varepsilon\in(0,\varepsilon_0)$, $\beta\in(0,\beta_0)$, $\theta\in(0,\theta_0)$ and $p\in(0,p_0)$.

Proof. For the sake of brevity we only prove the result for y near m_-. The proof for starting points near m_+ is similar. The eigenvalues of $\left(\frac{\partial^2}{\partial y_i \partial y_j} V(m_-)\right)_{i,j=1}^d$ are all positive. If $\|y-m_-\|$ is small we have

$$c\|y-m_-\|^2 \leqslant V(y) - V(m_-) \leqslant C\|y-m_-\|^2$$

for $C > c > 0$. Obviously it holds

$$\left\{\sup_{t\in[0,\varepsilon^{-\theta}]} \left\|\widetilde{Y}^\varepsilon_{r,t}(y) - m_-\right\| \geqslant \varepsilon^\beta\right\} \subseteq \left\{\sup_{t\in[0,\varepsilon^{-\theta}]} V(\widetilde{Y}^\varepsilon_{r,t}(y)) - V(m_-) \geqslant c\varepsilon^{2\beta}\right\}.$$

Choose N according to Lemma 4.13. The Itô formula gives

$$\begin{aligned} V(\widetilde{Y}^\varepsilon_{r,t}(y)) &= V(y) - \int_0^t \|\nabla V(\widetilde{Y}^\varepsilon_{r,s})\|^2 \,\mathrm{d}s + \varepsilon \int_0^t \nabla V(\widetilde{Y}^\varepsilon_{r,s}) \cdot \gamma(r+s)\,\mathrm{d}s \\ &\quad + \varepsilon \sum_{i=1}^d \int_r^{r+t} \frac{\partial}{\partial x_i} V(\widetilde{Y}^\varepsilon_{r,(s-r)-})\,\mathrm{d}\widetilde{A}^{M,i}_s \quad + \frac{\varepsilon^2}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2}{\partial y_i \partial y_j} V(\widetilde{Y}^\varepsilon_{r,s})\,\mathrm{d}\left[\widetilde{Y}^{\varepsilon,i}_{r,\cdot}, \widetilde{Y}^{\varepsilon,j}_{r,\cdot}\right]^c_s \\ &\quad + \sum_{s\leqslant t}\left(V(\widetilde{Y}^\varepsilon_{r,s}) - V(\widetilde{Y}^\varepsilon_{r,s-}) - \nabla V(\widetilde{Y}^\varepsilon_{r,s-})\Delta \widetilde{Y}^\varepsilon_{r,s}\right) \\ &=: I_1(t) + \ldots + I_6(t). \end{aligned}$$

Because of $\|y-m_-\| \leqslant \varepsilon^{2\beta}$ the inequality $V(y) - V(m_-) \leqslant C\varepsilon^4$ holds and verifies $\mathbb{P}(I_1(t) - V(m_-) > \varepsilon^{3\beta}) = 0$. As usual $I_2(t)$ is not positive. On the set $\left\{\sup_{t\leqslant\varepsilon^{-\theta}} \|\widetilde{Y}^\varepsilon_{r,t}(y)\| < N\right\}$ the boundedness of all involved integrands is satisfied. Additionally the estimate of

$$\sup_{r\geqslant 0} \sup_{y:\|y-m_-\|\leqslant\varepsilon^{2\beta}} \mathbb{P}\left(\left\{\sup_{t\in[0,\varepsilon^{-\theta}]} |I_i(t)| \geqslant \varepsilon^{3\beta}\right\} \cap \left\{\sup_{t\leqslant\varepsilon^{-\theta}} \|\widetilde{Y}^\varepsilon_{r,t}(y)\| < N\right\}\right)$$

for $i = 3, 4, 5$ exploits (A1), Lemma 4.11, the knowledge about the continuous part of the quadratic covariation of $\widetilde{Y}^\varepsilon_{r,t}$ and the local boundedness of the second order derivative of V. Again $I_6(t)$ is of order $\varepsilon^2 \sum_{s\leqslant t} \|\widetilde{A}^T_{r+s}\|^2$, which attributes this proof to Lemma 4.12. □

4.3.3 The exponential estimate

After the following statement we are well prepared to prove Theorem 4.10.

Lemma 4.16. *For $L \geqslant R^*$ and $\delta_0, \kappa > 0$ there exist ε_0, β_0 and $p_0 > 0$, such that*

$$\sup_{r\geqslant 0} \sup_{y\in\Omega_{\pm,L}(\delta_0)} \mathbb{P}\left(\sup_{t\in[0,\varepsilon^{-\kappa}]} \left\|\widetilde{Y}^\varepsilon_{r,t}(y) - y_t(y)\right\| \geqslant \varepsilon^\beta\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}}$$

holds for all $\varepsilon \in (0,\varepsilon_0), \beta \in (0,\beta_0)$ and $p \in (0,p_0)$.

Proof. **Step 1:** We claim the result with κ replaced by a sufficiently small parameter $\theta > 0$.

At first mimic the proof of Lemma 3.24 to get for all $r \geqslant 0$

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[0,\varepsilon^{-\theta}]}\left\|\widetilde{Y}^{\varepsilon}_{r,t}(y) - y_t(y)\right\| \geqslant \varepsilon^{\beta}\right) \\
&\leqslant \mathbb{P}\left(\sup_{t\in[0,R_{\varepsilon^{5\beta}}]}\left\|\widetilde{Y}^{\varepsilon}_{r,t}(y) - y_t(y)\right\| \geqslant \varepsilon^{5\beta}\right) \\
&\quad + \mathbb{P}\left(\sup_{t\in[0,\varepsilon^{-\theta}]}\left\|\widetilde{Y}^{\varepsilon}_{r,t}(y) - y_t(y)\right\| \geqslant \varepsilon^{\beta}, \sup_{t\in[0,R_{\varepsilon^{5\beta}}]}\left\|\widetilde{Y}^{\varepsilon}_{r,t}(y) - y_t(y)\right\| < \varepsilon^{5\beta}\right).
\end{aligned}$$

The first summand of the upper bound can be handled with Lemma 4.14. For the last one no excursion concerning periodic solutions is necessary. Simply estimate as below

$$\begin{aligned}
&\sup_{y\in\Omega_{\pm,L}(\delta_0)} \mathbb{P}\left(\sup_{t\in[0,\varepsilon^{-\theta}]}\left\|\widetilde{Y}^{\varepsilon}_{r,t}(y) - y_t(y)\right\| \geqslant \varepsilon^{\beta}, \sup_{t\in[0,R_{\varepsilon^{5\beta}}]}\left\|\widetilde{Y}^{\varepsilon}_{r,t}(y) - y_t(y)\right\| < \varepsilon^{5\beta}\right) \\
&\leqslant \sup_{s\geqslant 0}\ \sup_{y:\|y-m_\pm\|\leqslant\varepsilon^{4\beta}} \mathbb{P}\left(\sup_{t\in[0,\varepsilon^{-\theta}]}\left\|\widetilde{Y}^{\varepsilon}_{s,t}(y) - m_\pm\right\| \geqslant \varepsilon^{2\beta}\right)
\end{aligned}$$

and use Lemma 4.15 to guarantee the modified assertion of Lemma 4.16.

Step 2: Choose $\theta > 0$ small enough according to step one and $k_\varepsilon \in \mathbb{N}$ with $\varepsilon^{-\kappa} \leqslant k_\varepsilon \varepsilon^{-\theta}$. Define

$$B_j = \left\{\sup_{t\in[j\varepsilon^{-\theta},(j+1)\varepsilon^{-\theta}]}\left\|\widetilde{Y}^{\varepsilon}_{r,t}(y) - y_{t-j\varepsilon^{-\theta}}(\widetilde{Y}^{\varepsilon}_{r,j\varepsilon^{-\theta}}(y))\right\| < \varepsilon^{2\beta}\right\},$$

for $j \in \mathbb{N}_0$. Note that again the following inclusion is valid

$$\bigcap_{j=0}^{n-1} B_j \subseteq \left\{\sup_{t\in[0,n\varepsilon^{-\theta}]}\left\|\widetilde{Y}^{\varepsilon}_{r,t}(y) - y_t(y)\right\| < \varepsilon^{\beta}\right\}, \quad n \in \mathbb{N},$$

which is obvious for $n = 1$. Assume it holds for $n = k$ and now consider $n = k + 1$. From $R_{\varepsilon^{\beta}} < \varepsilon^{-\theta}$ and $\widetilde{Y}^{\varepsilon}_{r,(k-1)\varepsilon^{-\theta}}(y) \in \Omega_{\pm,L}(\delta_0)$ we immediately derive

$$\left\|y_{\varepsilon^{-\theta}}(\widetilde{Y}^{\varepsilon}_{r,(k-1)\varepsilon^{-\theta}}(y)) - m_\pm\right\| \leqslant \varepsilon^{2\beta}.$$

Hence on the event B_{k-1} the deviation $\|\widetilde{Y}^{\varepsilon}_{r,k\varepsilon^{-\theta}}(y) - m_\pm\|$ is smaller than $2\varepsilon^{2\beta}$. Putting together these results verifies for $t \in [k\varepsilon^{-\theta}, (k+1)\varepsilon^{-\theta}]$ on $\cap_{j=0}^{k} B_j$ that

$$\left\|\widetilde{Y}^{\varepsilon}_{r,t}(y) - y_t(y)\right\| < \varepsilon^{2\beta} + \left\|y_{t-k\varepsilon^{-\theta}}(\widetilde{Y}^{\varepsilon}_{r,k\varepsilon^{-\theta}}(y)) - y_t(y)\right\| < \varepsilon^{\beta}.$$

As earlier use the verified inclusion to get

$$\mathbb{P}\left(\sup_{t\in[0,\varepsilon^{-\kappa}]}\left\|\widetilde{Y}^{\varepsilon}_{r,t}(y) - y_t(y)\right\| \geqslant \varepsilon^{\beta}\right) \leqslant \mathbb{P}(B_0^c) + \sum_{j=1}^{k_\varepsilon-1}\mathbb{P}(B_0 \cap \ldots \cap B_{j-1} \cap B_j^c).$$

Consult the first step to deal with $\mathbb{P}(B_0^c)$ and again exploit the Markov property of the small jump process to estimate all the other summands as below

$$\begin{aligned}
&\mathbb{P}(B_0 \cap \ldots \cap B_{j-1} \cap B_j^c) \\
&\leqslant \mathbb{P}\Big(\widetilde{Y}^\varepsilon_{r,j\varepsilon^{-\theta}}(y) \in \Omega_{\pm,L}(\delta_0), \sup_{t\in[j\varepsilon^{-\theta},(j+1)\varepsilon^{-\theta}]} \big\|\widetilde{Y}^\varepsilon_{r,t}(y) - y_{t-j\varepsilon^{-\theta}}(\widetilde{Y}^\varepsilon_{r,j\varepsilon^{-\theta}}(y))\big\| \geqslant \varepsilon^{2\beta}\Big) \\
&\leqslant \sup_{r\geqslant 0} \sup_{y\in\Omega_{\pm,L}(\delta_0)} \mathbb{P}\left(\sup_{t\in[0,\varepsilon^{-\theta}]} \big\|\widetilde{Y}^\varepsilon_{r,t}(y) - y_t(y)\big\| \geqslant \varepsilon^{2\beta}\right).
\end{aligned}$$

Due to step one this upper bound is not bigger than $\mathrm{e}^{-\varepsilon^{-p}}$. Therefore the probability under consideration admits the upper bound $k_\varepsilon \mathrm{e}^{-\varepsilon^{-p}}$ which is smaller than $\mathrm{e}^{-\varepsilon^{-p/2}}$ for sufficiently small ε. □

Finally we prove Theorem 4.10.

Proof. Analogously to the proof of Theorem 3.18 it remains to estimate $\mathbb{P}(T_r \geqslant \varepsilon^{-\kappa})$. Direct calculation yields

$$\mathbb{P}(T_r \geqslant \varepsilon^{-\kappa}) = \exp\left(-\int_r^{r+\varepsilon^{-\kappa}} \beta^\varepsilon\left(\frac{t}{2T}\right) \mathrm{d}t\right).$$

But

$$\beta^\varepsilon\left(\frac{t}{2T}\right) = \varepsilon^{\rho\alpha(t/2T)}\nu_{t/2T}(B_1^c(0)) \geqslant \varepsilon^{\rho\alpha^*} \min_{t\in[0,1]} \nu_t(B_1^c(0))$$

implies

$$\mathbb{P}(T_r \geqslant \varepsilon^{-\kappa}) \leqslant \exp\left(-\varepsilon^{\rho\alpha^*-\kappa} \min_{t\in[0,1]} \nu_t(B_1^c(0))\right).$$

This upper bound is independent of $r > 0$ and tends to zero exponentially fast if $\kappa > \rho\alpha^*$. □

4.4 Exit times

In this section a ε-dependent time scale is revealed such that the scaled jump diffusion in the small noise limit shows a similar jump behaviour as a time-continuous two-state Markov chain that is only allowed to jump at time points $a + k - 1$, $k \in \mathbb{N}$, where $\alpha(t)$ is minimal. A usual tool to prove this is the analysis of appropriately scaled exit time of Y^ε from a region of attraction. The treatment of this exit time requires an estimate of the probability of a big deviation of the small jump process of Y^ε from the deterministic trajectory. Since we only know an upper bound of this probability for starting values in $\Omega_{\pm,L}(\delta_0)$ for some $\delta_0 > 0$ and $L \geqslant R^*$ (Theorem 4.10), we start to analyse exits from bounded and reduced domains and therefor fix a special error constant.

For all $\delta^* > 0$ there exist $L^* := L^*(\delta^*) > R^*$ and $\delta_1^* = \delta_1^*(\delta^*), \delta_2^* = \delta_2^*(\delta^*)$ with $0 < \delta_2^* < \delta_1^* < \delta^*$ such that the succeeding conditions hold

(D1) $\max_{t\in[0,1]} \nu_t\left((O_{L^*}^c - m_+) \cup (O_{L^*}^c - m_-)\right) < \frac{\delta^*}{4}$,

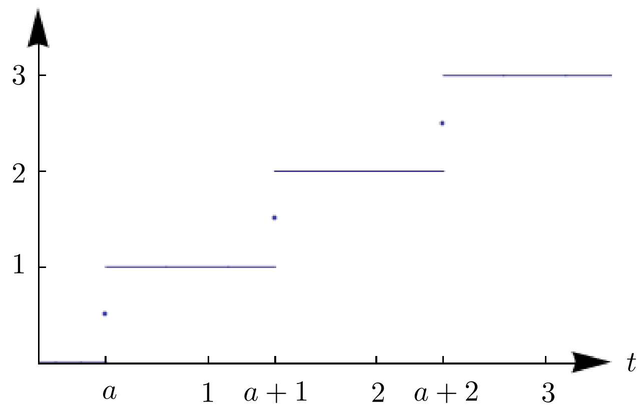


Figure 4.2: The step function g.

(D2) $\max_{t\in[0,1]} \nu_t\left((\Gamma_{L^*}(\delta_1^*,\delta_2^*,\delta_2^*)-m_+)\cup(\Gamma_{L^*}(\delta_1^*,\delta_2^*,\delta_2^*)-m_-)\right)<\frac{\delta^*}{4}$,

(D3) for all $L\geqslant L^*+\delta^* =: \hat{L}^*$: $O_L^{-2\delta_2^*}\supseteq O_{L^*}$ with $O_L^{-2\delta_2^*}=\{x\in O_L : \mathrm{dist}(\partial O_L,x)\geqslant 2\delta_2^*\}$.

The property (D3) is technical, not restrictive and we formulate it for convenience.
In the following $\delta^*,\delta_1^*,\delta_2^*$ and L^* are chosen such that (D1)-(D3) hold.

4.4.1 Exit from a bounded and reduced domain of attraction

Definition 4.17. *(i) Define* $2T=\lambda(\varepsilon)^{-1}$ *with* $\lambda(\varepsilon)=\varepsilon^{\alpha_*}|\log\varepsilon|^{-1/2}$ *and the parameters*

$$c_\pm=\sqrt{\frac{2\pi}{\alpha''(a)}}\,\nu_a(\Omega_\pm^c-m_\pm).$$

(ii) For $L\geqslant L^*+\delta^*$*, denote by* $\tau_\pm^{\varepsilon,*}$ *the exit time of* Y^ε *from the bounded and reduced domain* $\Omega_{\pm,L}(\delta_1^*)$*,*

$$\tau_\pm^{\varepsilon,*}=\inf\left\{t\geqslant 0 : Y_t^\varepsilon(y)\notin\Omega_{\pm,L}(\delta_1^*)\right\}.$$

(iii) Define the piecewise constant function g *(Figure 4.2) by*

$$g(t)=\begin{cases}0, & t<a,\\ k-\frac{1}{2}, & t=a+k-1,\ k\in\mathbb{N},\\ k, & t\in(a+k-1,a+k),\ k\in\mathbb{N}.\end{cases}$$

Theorem 4.18. *Assume* $y\in\Omega_\pm(\delta_1^*,\delta_2^*)\cap O_L^{-\delta_2^*}$ *and* $t>0$*. Then there is* $\varepsilon_0>0$ *such that for all* $\varepsilon<\varepsilon_0$

$$\left|\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\lambda(\varepsilon)>t\right)-\mathrm{e}^{-g(t)c_\pm}\right|\leqslant\delta^*.$$

Remark 4.19. The estimate in the given theorem holds pointwise in t and is not uniform.

The following lemma and Laplace's method will suffice to verify Theorem 4.18.

Lemma 4.20. *Let $t > 0$ and for $L \geqslant L^* + \delta^*$ choose $y \in \Omega_\pm(\delta_1^*, \delta_2^*) \cap O_L^{-\delta_2^*}$. Then there is $\varepsilon_0 > 0$ such that*

$$\left|\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\lambda(\varepsilon) \leqslant t\right) - I^\varepsilon(t)\right| \leqslant \delta^*$$

for $\varepsilon < \varepsilon_0$, where

$$I^\varepsilon(t) = 1 - \exp\left(-\int_0^t \sqrt{|\log\varepsilon|}\varepsilon^{\alpha(r)-\alpha_*}\nu_r(\Omega_{\pm,L}^c - m_\pm)\,\mathrm{d}r\right).$$

The main part of this subsection is devoted to the proof of Lemma 4.20. But at first some important definitions and preparatory results are necessary.

Definition 4.21. *Assume $y \in \mathbb{R}^d$, $L \geqslant L^* + \delta^*$ and $0 \leqslant s \leqslant t$. The random variable W_t^ε is distributed according to the law $\nu_{t/2T}\left(B_{\varepsilon^{-\rho}}^c(0) \cap \frac{1}{\varepsilon}\cdot\right)/\nu_{t/2T}(B_{\varepsilon^{-\rho}}^c(0))$. Define*

$$\begin{aligned}
B_{s,t}^\pm(y) &= \left\{\widetilde{Y}_{s,r}^\varepsilon(y) \in \Omega_{\pm,L}(\delta_1^*), r \in [0, t-s), \widetilde{Y}_{s,t-s}^\varepsilon(y) + W_t^\varepsilon \in \Omega_{\pm,L}(\delta_1^*)\right\},\\
\bar{B}_{s,t}^\pm(y) &= \left\{\widetilde{Y}_{s,r}^\varepsilon(y) \in \Omega_{\pm,L}(\delta_1^*), r \in [0, t-s), \widetilde{Y}_{s,t-s}^\varepsilon(y) + W_t^\varepsilon \in \Omega_\pm(\delta_1^*, \delta_2^*) \cap O_L^{-\delta_2^*}\right\},\\
\hat{B}_{s,t}^\pm(y) &= \Big\{\widetilde{Y}_{s,r}^\varepsilon(y) \in \Omega_{\pm,L}(\delta_1^*), r \in [0, t-s),\\
&\qquad \widetilde{Y}_{s,t-s}^\varepsilon(y) + W_t^\varepsilon \in \Omega_{\pm,L}(\delta_1^*)\setminus\left(\Omega_\pm(\delta_1^*, \delta_2^*) \cap O_L^{-\delta_2^*}\right)\Big\},\\
C_{s,t}^\pm(y) &= \left\{\widetilde{Y}_{s,r}^\varepsilon(y) \in \Omega_{\pm,L}(\delta_1^*), r \in [0, t-s), \widetilde{Y}_{s,t-s}^\varepsilon(y) + W_t^\varepsilon \notin \Omega_{\pm,L}(\delta_1^*)\right\},\\
E_{s,t}(y) &= \left\{\sup_{r\in[0,t-s]} \|\widetilde{Y}_{s,r}^\varepsilon(y) - y_r(y)\| \leqslant \frac{1}{2}\varepsilon^{2\gamma}\right\},\\
F_{s,t}^\pm(y) &= \left\{\exists u \in (0, t-s) : \widetilde{Y}_{s,r}^\varepsilon \in \Omega_{\pm,L}(\delta_1^*), r \in [0, u), \widetilde{Y}_{s,u}^\varepsilon \notin \Omega_{\pm,L}(\delta_1^*)\right\}.
\end{aligned}$$

Lemma 4.22. *Assume $k \in \mathbb{N}\setminus\{1\}$, $t > 0$ and $B_0, \ldots, B_k$ equal to $\Omega_{\pm,L}(\delta_1^*)$, $\Omega_\pm(\delta_1^*, \delta_2^*) \cap O_L^{-\delta_2^*}$, $\Omega_{\pm,L}(\delta_1^*)\setminus\left(\Omega_\pm(\delta_1^*, \delta_2^*) \cap O_L^{-\delta_2^*}\right)$ or $(\Omega_{\pm,L}(\delta_1^*))^c$. For all $u \geqslant 0$ the random variable W_u^ε is distributed according to the law $\nu_{u/2T}\left(B_{\varepsilon^{-\rho}}^c(0) \cap \frac{1}{\varepsilon}\cdot\right)/\nu_{u/2T}(B_{\varepsilon^{-\rho}}^c(0))$. For all $j = 0, \ldots, k-2$, $0 \leqslant s \leqslant r$ and $y \in \mathbb{R}^d$ the set $S_{s,r}^j(y)$ equals to*

$$S_{s,r}^j(y) = \left\{\widetilde{Y}_{s,v}^\varepsilon(y) \in B_j, v \in [0, r-s), \widetilde{Y}_{s,r-s}^\varepsilon(y) + W_r^\varepsilon \in B_{j+1}\right\}.$$

Define $s_0 = 0$ and choose $y_0 \in B_0$. Recall $2T = \varepsilon^{-\alpha_*}\sqrt{|\log\varepsilon|}$. Then the following estimate holds

$$\begin{aligned}
&\mathbb{E}\,\mathbf{1}_{\{\tau_k\leqslant 2Tt\}}\prod_{j=0}^{k-1}\mathbf{1}\left(S_{\tau_j,\tau_{j+1}}^j(Y_{\tau_j}^\varepsilon(y_0))\right)\\
&\geqslant \int_0^t\int_0^{s_k}\cdots\int_0^{s_2}\left[\prod_{j=1}^k\left(\varepsilon^{\rho\alpha(s_j)-\alpha_*}\sqrt{|\log\varepsilon|}\nu_{s_j}(B_1^c(0))\right)\mathrm{e}^{-\int_0^{s_k}\varepsilon^{\rho\alpha(r)-\alpha_*}\sqrt{|\log\varepsilon|}\nu_r(B_1^c(0))\mathrm{d}r}\right.\\
&\left.\prod_{j=0}^{k-1}\inf_{y\in B_j}\mathbb{P}(S_{2Ts_j,2Ts_{j+1}}^j(y))\right]\mathrm{d}s_1\ldots\mathrm{d}s_k.
\end{aligned}$$

A similar upper bound exists with suprema instead of infima.

Proof. Pursue the same strategy as in the proof of Lemma 3.30. Repeatedly use the strong Markov property of $(Y_t^\varepsilon, t)_{t\geqslant 0}$ (Proposition 2.45) and afterwards the knowledge of the conditional density of the inter-jump times (Lemma 4.8) to get

$$\begin{aligned}
&\mathbb{E}\,\mathbf{1}_{\{\tau_k\leqslant 2Tt\}}\prod_{j=0}^{k-1}\mathbf{1}\left(S^j_{\tau_j,\tau_{j+1}}(Y^\varepsilon_{\tau_j}(y_0))\right)\\
&\geqslant \int_0^\infty\cdots\int_0^\infty\Bigg[\prod_{j=1}^{k}\left(\varepsilon^{\rho\alpha((t_1+\ldots+t_j)/2T)}\nu_{(t_1+\ldots+t_j)/2T}(B_1^c(0))\right)\mathrm{e}^{-\int_0^{t_1+\ldots+t_k}\varepsilon^{\rho\alpha(r/2T)}\nu_{r/2T}(B_1^c(0))\mathrm{d}r}\\
&\quad\mathbf{1}_{[0,t]}\left(\frac{t_1+\ldots+t_k}{2T}\right)\prod_{j=0}^{k-1}\inf_{y\in B_j}\mathbb{P}(S^j_{t_0+\ldots+t_j,t_1+\ldots+t_{j+1}}(y))\Bigg]\,\mathrm{d}t_k\ldots\mathrm{d}t_1
\end{aligned}$$

with $t_0=0$. The substitutions $(s_1,s_2\ldots,s_k)=\left(\frac{t_1}{2T},\frac{t_1+t_2}{2T},\ldots,\frac{t_1+\ldots+t_k}{2T}\right)$ and $u=\frac{r}{2T}$ in the exponent verify the lemma. ☐

Remember the δ-enlargement $B^{+\delta}=B\cup\left\{x\in\mathbb{R}^d:\operatorname{dist}(x,B)\leqslant\delta\right\}$ of a Borel set B.

Lemma 4.23. *a) Let $L\geqslant L^*$ and $\delta>0$ then there is $L_\delta(L)>L$ such that $O_L^{+\delta}\subseteq O_{L_\delta(L)}$.*

b) For all $L\geqslant L^*+\delta^*$ we have

- *(i)* $(\Omega_{\pm,L}(\delta_1^*))^{+\delta_2^*}\subseteq\Omega_{\pm,L_{\delta^*}(L)}$,
- *(ii)* $(\Omega^c_{\pm,L}(\delta_1^*))^{+\delta_2^*}\subseteq\Omega^c_{\pm,L^*}(\delta_1^*,\delta_2^*)$,
- *(iii)* $(\Omega^c_{\pm,L_{\delta^*}(L)})^{+\delta_2^*}\subseteq\Omega^c_{\pm,L}(\delta_1^*)$,
- *(iv)* $(\Omega_{\pm,L^*}(\delta_1^*,\delta_2^*,\delta_2^*))^{+\delta_2^*}\subseteq\Omega_\pm(\delta_1^*,\delta_2^*)\cap O_L^{-\delta_2^*}$,
- *(v)* $\left(\Omega_{\pm,L}(\delta_1^*)\backslash\left(\Omega_\pm(\delta_1^*,\delta_2^*)\cap O_L^{-\delta_2^*}\right)\right)^{+\delta_2^*}\subseteq\Gamma_{L^*}(\delta_1^*,\delta_2^*,\delta_2^*)\cup O^c_{L^*}$.

Proof. The first statement is a consequence of the Lipschitz continuity of V on compacts. Define $L_\delta(L)\coloneqq L+\delta C_L$ if C_L denotes the Lipschitz constant of V on $O_L^{+\delta}$.
The proof of assertion (i) falls back on $(\Omega_\pm(\delta_1^*))^{+\delta_2^*}\subseteq\Omega_\pm$, because $\delta_1^*>\delta_2^*$ is true and the inclusion $O_L^{+\delta_2^*}\subseteq O_{L_{\delta^*}(L)}$ holds since $\delta_2^*<\delta^*$. Passing to the complements of all sets involved in (i) immediately proves (iii). From (D3) we can deduce $(O_L^{-\delta_2^*})^c\subseteq O^c_{L^*}$. In addition $(\Omega^c_\pm(\delta_1^*))^{+\delta_2^*}\subseteq\Omega^c_\pm(\delta_1^*,\delta_2^*)$ yields (ii). The proof of (iv) demands $(\Omega_\pm(\delta_1^*,\delta_2^*,\delta_2^*))^{+\delta_2^*}\subseteq\Omega_\pm(\delta_1^*,\delta_2^*)$. The inclusion in (v) is a consequence of $O_L^{-2\delta_2^*}\supseteq O_{L_*}$ for $L\geqslant L^*+\delta^*$ and

$$\left(\Omega_{\pm,L}(\delta_1^*)\backslash\left(\Omega_\pm(\delta_1^*,\delta_2^*)\cap O_L^{-\delta_2^*}\right)\right)^{+\delta_2^*}\subseteq\left(O_L\backslash O_L^{-\delta_2^*}\right)^{+\delta_2^*}\cup(\Gamma_L(\delta_1^*,\delta_2^*))^{+\delta_2^*}.$$

☐

Lemma 4.24. *Assume $\rho\in(0,1)$. Let $C>0$ and $L\geqslant L^*+\delta^*$. Define $L_\delta(L)$ as in the previous lemma. There exist $\varepsilon_0,p_0,\gamma_0>0$ such that the succeeding inequalities hold for all $s,t\geqslant 0$ with $t\leqslant C$, $\gamma\in(0,\gamma_0)$, $p\in(0,p_0)$ and $\varepsilon\in(0,\varepsilon_0)$:*

(i)

$$\inf_{y\in\Omega_\pm(\delta_1^*,\delta_2^*)\cap O_L^{-\delta_2^*}} \mathbb{P}(\bar{B}^\pm_{s2T,(s+t)2T}(y))\mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}$$
$$\geqslant \mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}\left(1-\mathrm{e}^{-\varepsilon^{-p}}\right)\frac{\nu_{s+t}\left(B^c_{\varepsilon^{-\rho}}(0)\cap\frac{1}{\varepsilon}\left(\Omega_{\pm,L^*}(\delta_1^*,\delta_2^*,\delta_2^*)-m_\pm\right)\right)}{\varepsilon^{\rho\alpha(s+t)}\nu_{s+t}(B_1^c(0))},$$

(ii)

$$\inf_{y\in\Omega_\pm(\delta_1^*,\delta_2^*)\cap O_L^{-\delta_2^*}} \mathbb{P}(C^\pm_{s2T,(s+t)2T}(y))\mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}$$
$$\geqslant \mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}\left(1-\mathrm{e}^{-\varepsilon^{-p}}\right)\frac{\nu_{s+t}\left(\frac{1}{\varepsilon}\left(\Omega^c_{\pm,L_{\delta^*}(L)}-m_\pm\right)\right)}{\varepsilon^{\rho\alpha(s+t)}\nu_{s+t}(B_1^c(0))},$$

(iii)

$$\sup_{y\in\Omega_{\pm,L}(\delta_1^*)} \mathbb{P}(B^\pm_{s2T,(s+t)2T}(y))\mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}\leqslant \mathrm{e}^{-\varepsilon^{-p}}+\frac{\nu_{s+t}\left(B^c_{\varepsilon^{-\rho}}(0)\cap\frac{1}{\varepsilon}\left(\Omega_{\pm,L_{\delta^*}(L)}-m_\pm\right)\right)}{\varepsilon^{\rho\alpha(s+t)}\nu_{s+t}(B_1^c(0))},$$

(iv)

$$\sup_{y\in\Omega_{\pm,L}(\delta_1^*)} \mathbb{P}(C^\pm_{s2T,(s+t)2T}(y))\mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}\leqslant \mathrm{e}^{-\varepsilon^{-p}}+\frac{\nu_{s+t}\left(\frac{1}{\varepsilon}\left(\Omega^c_{\pm,L^*}(\delta_1^*,\delta_2^*)-m_\pm\right)\right)}{\varepsilon^{\rho\alpha(s+t)}\nu_{s+t}(B_1^c(0))},$$

(v)

$$\sup_{y\in\Omega_{\pm,L}(\delta_1^*)} \mathbb{P}(\hat{B}^\pm_{s2T,(s+t)2T}(y))\mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}\leqslant \mathrm{e}^{-\varepsilon^{-p}}+\frac{\nu_{s+t}\left(\frac{1}{\varepsilon}\left(\Gamma_{L^*}(\delta_1^*,\delta_2^*,\delta_2^*)\cup O^c_{L^*}\right)\right)}{\varepsilon^{\rho\alpha(s+t)}\nu_{s+t}(B_1^c(0))}.$$

Proof. Mainly we will follow the arguments of the proof of Lemma 3.31. Intersect with $E_{s2T,(s+t)2T}(y)$, choose ε small enough such that $\frac{1}{2}\varepsilon^{2\gamma}<\delta_2^*$ holds, use the Markov property of Y^ε and apply Lemma 4.23 (iv) to get

$$\mathbb{P}\left(\bar{B}^\pm_{s2T,(s+t)2T}(y)\right)\mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}$$
$$\geqslant \mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}\mathbb{P}\left(E_{s2T,(s+t)2T}(y)\right)\mathbb{P}\left(W^\varepsilon_{(s+t)2T}\in\Omega_{\pm,L^*}(\delta_1^*,\delta_2^*,\delta_2^*)-m_\pm\right).$$

Recall that W_u^ε used in the definition of $B^\pm_{v,u}(y)$ is distributed according to the law $\nu_{u/2T}\left(B^c_{\varepsilon^{-\rho}}(0)\cap\frac{1}{\varepsilon}\cdot\right)/\nu_{u/2T}(B^c_{\varepsilon^{-\rho}}(0))$. Insert this law and estimate $\mathbb{P}(E_{s2T,(s+t)2T}(y))$ by Lemma 4.16 since $t2T$ is at most of polynomial order in ε due to $t\leqslant C$. This verifies (i). Assertion (ii) falls back on the same steps and exploits Lemma 4.23 (iii). For the given upper estimates again intersect with a nice small jump process and a very rough behaviour. Analogously arguments as in Lemma 3.31 together with an application of Lemma 4.23 (i), (ii) and (v) will complete this proof. □

It is time to reap the benefits of all these preparations. First prove Lemma 4.20 and subsequently verify Theorem 4.18.

Proof of the Lemma 4.20:

Proof. **Lower bound:**
For $\rho \in (\frac{2}{3}, 1)$ the inequality $\frac{1}{2}\alpha_*\rho > \alpha_*(1-\rho)$ is true. We define $k_\varepsilon = \lfloor \varepsilon^{-r} \rfloor$ for some $r \in (\alpha_*(1-\rho), \frac{1}{2}\alpha_*\rho)$ and for $y \in \Omega_\pm(\delta_1^*, \delta_2^*) \cap O_L^{-\delta_2^*}$ we estimate

$$\begin{aligned}\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\lambda(\varepsilon) \leqslant t\right) &\geqslant \sum_{k=1}^{k_\varepsilon} \mathbb{P}_y\left(\tau_\pm^{\varepsilon,*} = \tau_k, \tau_k\lambda(\varepsilon) \leqslant t\right) \\ &\geqslant \sum_{k=1}^{k_\varepsilon} \mathbb{E}_y\left[\mathbf{1}_{\{\tau_k\lambda(\varepsilon)\leqslant t\}} \prod_{j=0}^{k-2} \mathbf{1}\left(\bar{B}^\pm_{\tau_j,\tau_{j+1}}(Y^\varepsilon_{\tau_j})\right)\mathbf{1}\left(C^\pm_{\tau_{k-1},\tau_k}(Y^\varepsilon_{\tau_{k-1}})\right)\right].\end{aligned}$$

Now use Lemma 4.22 with $S^j_{\tau_j,\tau_{j+1}}(y) = \bar{B}^\pm_{\tau_j,\tau_{j+1}}(y)$ for $j = 0,\ldots,k-2$ and $S^{k-1}_{\tau_{k-1},\tau_k}(y) = C^\pm_{\tau_{k-1},\tau_k}(y)$ and $2T = \lambda(\varepsilon)^{-1}$ to get

$$\begin{aligned}\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\lambda(\varepsilon) \leqslant t\right) \geqslant &\sum_{k=1}^{k_\varepsilon} \int_0^t \int_0^{s_k} \cdots \int_0^{s_2} \Bigg[\prod_{j=1}^{k} \sqrt{|\log\varepsilon|}\varepsilon^{\rho\alpha(s_j)-\alpha_*}\nu_{s_j}(B_1^c(0)) \\ &\mathrm{e}^{-\int_0^{s_k}\sqrt{|\log\varepsilon|}\varepsilon^{\rho\alpha(r)-\alpha_*}\nu_r(B_1^c(0))\mathrm{d}r} \inf_{y\in\Omega_\pm(\delta_1^*,\delta_2^*)\cap O_L^{-\delta_2^*}} \mathbb{P}(C^\pm_{s_{k-1}2T,s_k2T}(y)) \\ &\prod_{j=0}^{k-2} \inf_{y\in\Omega_\pm(\delta_1^*,\delta_2^*)\cap O_L^{-\delta_2^*}} \mathbb{P}(\bar{B}^\pm_{s_j2T,s_{j+1}2T}(y))\Bigg]\mathrm{d}s_1\ldots\mathrm{d}s_k.\end{aligned}$$

Introduce the factor $\prod_{j=0}^{k-1}\mathbf{1}_{\{(s_{j+1}-s_j)2T\geqslant R_\varepsilon\gamma\}} \in [0,1]$ and use Lemma 4.24 (i) and (ii) to prove

$$\begin{aligned}&\mathbb{P}_y(\tau_\pm^{\varepsilon,*}\lambda(\varepsilon) \leqslant t) \\ &\geqslant \sum_{k=1}^{k_\varepsilon} \int_0^t \int_0^{s_k} \cdots \int_0^{s_2} \Bigg[\left(\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\right)^k \mathrm{e}^{-\int_0^{s_k}\sqrt{|\log\varepsilon|}\varepsilon^{\rho\alpha(r)-\alpha_*}\nu_r(B_1^c(0))\mathrm{d}r} \\ &\prod_{j=0}^{k-1}\mathbf{1}_{\{(s_{j+1}-s_j)2T\geqslant R_\varepsilon\gamma\}} \prod_{j=1}^{k-1}\nu_{s_j}\left(B^c_{\varepsilon^{-\rho}}(0) \cap \frac{1}{\varepsilon}\left(\Omega_{\pm,L^*}(\delta_1^*,\delta_2^*,\delta_2^*) - m_\pm\right)\right) \\ &\left(1-\mathrm{e}^{-\varepsilon^{-p}}\right)^k \nu_{s_k}\left(\frac{1}{\varepsilon}\left(\Omega^c_{\pm,L_{\delta^*}(L)} - m_\pm\right)\right)\Bigg]\mathrm{d}s_1\ldots\mathrm{d}s_k.\end{aligned} \tag{4.4}$$

We will get rid of the factor concerning the logarithmic return times through the following estimate

$$\prod_{j=0}^{k-1}\mathbf{1}_{\{(s_{j+1}-s_j)2T\geqslant R_\varepsilon\gamma\}} \geqslant 1 - \sum_{j=0}^{k-1}\mathbf{1}_{\{(s_{j+1}-s_j)2T< R_\varepsilon\gamma\}}.$$

But at first create another additive error term through adding the missing summands $k > k_\varepsilon$

of a useful series representation. This is all done in the following:

$$\begin{aligned}
&\mathbb{P}_y(\tau_{\pm}^{\varepsilon,*}\lambda(\varepsilon) \leqslant t) \qquad\qquad (4.5)\\
&\geqslant \sum_{k=1}^{\infty} \int_0^t \int_0^{s_k} \cdots \int_0^{s_2} \Bigg[\left(\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\right)^k \mathrm{e}^{-\int_0^{s_k} \sqrt{|\log\varepsilon|}\varepsilon^{\rho\alpha(r)-\alpha_*}\nu_r(B_1^c(0))\mathrm{d}r} \left(1-\mathrm{e}^{-\varepsilon^{-p}}\right)^k \\
&\prod_{j=1}^{k-1} \nu_{s_j}\left(B_{\varepsilon^{-\rho}}^c(0) \cap \frac{1}{\varepsilon}\left(\Omega_{\pm,L^*}(\delta_1^*,\delta_2^*,\delta_2^*) - m_{\pm}\right)\right) \nu_{s_k}\left(\frac{1}{\varepsilon}\left(\Omega_{\pm,L_{\delta^*}(L)}^c - m_{\pm}\right)\right)\Bigg] \mathrm{d}s_1\ldots\mathrm{d}s_k \\
&\quad - \sum_{k=1}^{k_\varepsilon}\sum_{j=0}^{k-1} \int_0^\infty \int_0^{s_k} \cdots \int_0^{s_2} \Bigg[\left(\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\right)^k \prod_{i=1}^k \beta^\varepsilon(s_i)\, \mathrm{e}^{-\int_0^{s_k}\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\beta^\varepsilon(r)\mathrm{d}r} \\
&\quad\ \mathbf{1}_{\{(s_{j+1}-s_j)2T<R_\varepsilon\gamma\}}\Bigg] \mathrm{d}s_1\ldots\mathrm{d}s_k \\
&\quad - \sum_{k=k_\varepsilon+1}^{\infty} \int_0^t \int_0^{s_k} \cdots \int_0^{s_2} \prod_{j=1}^k \left(\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\beta^\varepsilon(s_j)\right) \mathrm{e}^{-\int_0^{s_k}\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\beta^\varepsilon(r)\mathrm{d}r} \mathrm{d}s_1\ldots\mathrm{d}s_k.
\end{aligned}$$

The first infinite sum in (4.5) contributes the main part to the probability and two error terms are substracted. The first one is of order $|\log\varepsilon|\varepsilon^{\rho\alpha_*-2r}$ what is explained in the following. The substitution $(t_1,t_2,\ldots,t_k) = (s_1, s_2-s_1,\ldots,s_k-s_{k-1})$ and $\int_0^\infty f(s+t)\mathrm{e}^{-\int_s^{s+t} f(r)\mathrm{d}r}\mathrm{d}t = 1$ for nonnegative functions f with $\int_s^\infty f(r)\mathrm{d}r = \infty$, $s \geqslant 0$, yields for $k = 1,\ldots,k_\varepsilon$ and $j = 0,\ldots,k-1$

$$\begin{aligned}
&\int_0^\infty \int_0^{s_k} \cdots \int_0^{s_2} \prod_{i=1}^k \left(\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\beta^\varepsilon(s_i)\right) \mathrm{e}^{-\int_0^{s_k}\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\beta^\varepsilon(r)\mathrm{d}r} \mathbf{1}_{\{(s_{j+1}-s_j)2T<R_\varepsilon\gamma\}}\mathrm{d}s_1\ldots\mathrm{d}s_k \\
&= \int_0^\infty \cdots \int_0^\infty \Bigg[\prod_{i=1}^k \left(\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\beta^\varepsilon(t_1+\ldots+t_i)\, \mathrm{e}^{-\int_{t_0+\ldots+t_{i-1}}^{t_1+\ldots+t_i}\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\beta^\varepsilon(r)\mathrm{d}r}\right) \\
&\quad \mathbf{1}_{\{t_{j+1}<R_\varepsilon\gamma/2T\}}\Bigg] \mathrm{d}t_k\ldots\mathrm{d}t_1 \qquad\qquad (4.6)\\
&= \int_0^\infty \cdots \int_0^\infty \Bigg[\prod_{i=1}^j \left(\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\beta^\varepsilon(t_1+\ldots+t_i)\, \mathrm{e}^{-\int_{t_0+\ldots+t_{i-1}}^{t_1+\ldots+t_i}\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\beta^\varepsilon(r)\mathrm{d}r}\right) \\
&\quad \left(1-\exp\left\{-\int_{t_0+\ldots+t_j}^{t_1+\ldots+t_j+(R_\varepsilon\gamma/2T)} \sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\beta^\varepsilon(r)\mathrm{d}r\right\}\right)\Bigg] \mathrm{d}t_j\ldots\mathrm{d}t_1 \\
&\leqslant \gamma C \max_{t\in[0,1]} \nu_t(B_1^c(0))\varepsilon^{\rho\alpha_*}|\log\varepsilon|,
\end{aligned}$$

with $R_\varepsilon\gamma = \gamma C|\log\varepsilon|$. This error occurs at most k_ε^2-times. Thus r must be smaller than $\frac{\rho\alpha_*}{2}$. Extending the finite sum of formula (4.4) to an infinite one creates an additive error (last sum in formula (4.5)) that is bounded from above by $\mathbb{P}\left(N^{A^T}\left(B_{\varepsilon^{-\rho}}^c(0)\times[0,t\lambda(\varepsilon)^{-1}]\right) > k_\varepsilon\right)$. Recall that the number of big jumps until time $t\lambda(\varepsilon)^{-1}$ is distributed according to the Poissonian law. Lemma 3.33 implies

$$\begin{aligned}
&\mathbb{P}\left(N^{A^T}\left(B^c_{\varepsilon^{-\rho}}(0)\times[0,t\lambda(\varepsilon)^{-1}]\right)>k_\varepsilon\right)\\
&=\sum_{k=k_\varepsilon+1}^{\infty}\mathrm{e}^{-\int_0^{t\lambda(\varepsilon)^{-1}}\beta^\varepsilon\left(\frac{s}{2T}\right)\mathrm{d}s}\frac{1}{k!}\left(\int_0^{t\lambda(\varepsilon)^{-1}}\beta^\varepsilon\left(\frac{s}{2T}\right)\mathrm{d}s\right)^k\\
&\leqslant\frac{1}{(k_\varepsilon+1)!}\left(\int_0^{t\lambda(\varepsilon)^{-1}}\beta^\varepsilon\left(\frac{s}{2T}\right)\mathrm{d}s\right)^{k_\varepsilon+1}.
\end{aligned}$$

From $\beta^\varepsilon(s)\leqslant\varepsilon^{\rho\alpha_*}\max_{t\in[0,1]}\nu_t(B^c_1(0))$ for all $s\geqslant 0$ we can derive

$$\mathbb{P}\left(N^{A^T}\left(B^c_{\varepsilon^{-\rho}}(0)\times[0,t\lambda(\varepsilon)^{-1}]\right)>k_\varepsilon\right)\leqslant c_1t\varepsilon^{\rho\alpha_*-\alpha_*}\sqrt{|\log\varepsilon|}\left(tc_2\sqrt{|\log\varepsilon|}\varepsilon^{\rho\alpha_*-\alpha_*}k_\varepsilon^{-1}\right)^{k_\varepsilon},$$

for some $c_1,c_2>0$ because $(k_\varepsilon!)^{-1}\leqslant(2\pi k_\varepsilon)^{-1/2}\,k_\varepsilon^{-k_\varepsilon}\mathrm{e}^{k_\varepsilon}$ due to Stirling's formula. Thus this error is a null sequence as $\varepsilon\to 0$ if $r>\alpha_*(1-\rho)$. It is left to analyse the first series of formula (4.5). Apply Lemma 3.32 (ii) and $\nu_t(\frac{1}{\varepsilon}A)=\varepsilon^{\alpha(t)}\nu_t(A)$ to transform the main part of the probability as below

$$\begin{aligned}
&\sum_{k=1}^{\infty}\int_0^t\int_0^{s_k}\cdots\int_0^{s_2}\left(\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\right)^k e^{-\int_0^{s_k}\sqrt{|\log\varepsilon|}\varepsilon^{\rho\alpha(r)-\alpha_*}\nu_r(B^c_1(0))\mathrm{d}r}\left(1-\mathrm{e}^{-\varepsilon^{-p}}\right)^k\\
&\prod_{j=1}^{k-1}\nu_{s_j}\left(B^c_{\varepsilon^{-\rho}}(0)\cap\frac{1}{\varepsilon}\left(\Omega_{\pm,L^*}(\delta_1^*,\delta_2^*,\delta_2^*)-m_\pm\right)\right)\nu_{s_k}\left(\frac{1}{\varepsilon}\left(\Omega^c_{\pm,L_{\delta^*}(L)}-m_\pm\right)\right)\mathrm{d}s_1\ldots\mathrm{d}s_k\\
&=\int_0^t\sqrt{|\log\varepsilon|}\varepsilon^{\alpha(s)-\alpha_*}\left(1-\mathrm{e}^{-\varepsilon^{-p}}\right)\nu_s\left(\Omega^c_{\pm,L_{\delta^*}(L)}-m_\pm\right)\\
&\qquad\exp\left\{-\int_0^s\sqrt{|\log\varepsilon|}\varepsilon^{\alpha(r)-\alpha_*}\left(1-\mathrm{e}^{-\varepsilon^{-p}}\right)\nu_r\left(\Omega^c_{\pm,L^*}(\delta_1^*,\delta_2^*,\delta_2^*)-m_\pm\right)\mathrm{d}r\right\}\mathrm{d}s.
\end{aligned}$$

As usual ignoring the terms $\mathrm{e}^{-\varepsilon^{-p}}$ produces only small errors. Remember the choice of the constants δ_1^*,δ_2^* and L_* which guarantees that the Lévy measures ν_t only have little weight near the separatrix and far away from the minima. Thus

$$\begin{aligned}
|\nu_t(\Omega^c_{\pm,L_{\delta^*}(L)}-m_\pm)-\nu_t(\Omega^c_{\pm,L}-m_\pm)|&\leqslant\nu_t(O^c_{L^*}-m_\pm)\leqslant\delta^*,\\
|\nu_r\left(\Omega^c_{\pm,L^*}(\delta_1^*,\delta_2^*,\delta_2^*)-m_\pm\right)-\nu_r(\Omega^c_{\pm,L}-m_\pm)|&\leqslant\nu_t(O^c_{L^*}-m_\pm)+\nu_t(\Gamma_{L^*}(\delta_1^*,\delta_2^*,\delta_2^*))\leqslant\delta^*.
\end{aligned}$$

From replacing $\nu_s(\Omega^c_{\pm,L_{\delta^*}(L)}-m_\pm)$ by $\nu_s(\Omega^c_{\pm,L}-m_\pm)-\delta^*$ and $\nu_r(\Omega^c_{\pm,L^*}(\delta_1^*,\delta_2^*,\delta_2^*)-m_\pm)$ by $\nu_r(\Omega^c_{\pm,L}-m_\pm)+\delta^*$ in the main part of the probability given above we can derive

$$\begin{aligned}
&\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\lambda(\varepsilon)\leqslant t\right)\\
&\geqslant\int_0^t\sqrt{|\log\varepsilon|}\varepsilon^{\alpha(s)-\alpha_*}(\nu_s(\Omega^c_{\pm,L}-m_\pm)-\delta^*)\,\mathrm{e}^{-\int_0^s\sqrt{|\log\varepsilon|}\varepsilon^{\alpha(r)-\alpha_*}(\nu_r(\Omega^c_{\pm,L}-m_\pm)+\delta^*)\mathrm{d}r}\mathrm{d}s-\mathrm{o}_\varepsilon(1).
\end{aligned}$$

The error term $\mathrm{o}_\varepsilon(1)$ is a null sequence as $\varepsilon\to 0$. It is of order ε^k for example with $k=\frac{1}{2}((\alpha_*\rho-2r)\wedge(\alpha_*\rho-\alpha_*+r))$. For sufficiently small ε and $c=2\left(\min_{s\in[0,1]}\nu_s(\Omega^c_{\pm,L}-m_\pm)\right)^{-1}$

we get

$$\begin{aligned}\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\lambda(\varepsilon)\leqslant t\right) &\geqslant -\delta^* + (1-c\delta^*)\left(1-\mathrm{e}^{-\int_0^t \sqrt{|\log\varepsilon|}\varepsilon^{\alpha(s)-\alpha_*}\nu_s(\Omega_{\pm,L}^c-m_\pm)\mathrm{d}s}\right)\\ &\geqslant -\delta^*(1+c) + \left(1-\mathrm{e}^{-\int_0^t \sqrt{|\log\varepsilon|}\varepsilon^{\alpha(s)-\alpha_*}\nu_s(\Omega_{\pm,L}^c-m_\pm)\mathrm{d}s}\right)\end{aligned}$$

which proves the result.

Upper bound: As before define $k_\varepsilon = \lfloor \varepsilon^{-r}\rfloor$ for some $r \in (\alpha_*(1-\rho), \frac{\alpha_*\rho}{2})$, with $\rho\in(\frac{2}{3},1)$ and estimate

$$\begin{aligned}\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\lambda(\varepsilon)\leqslant t\right) &\leqslant \sum_{k=1}^{k_\varepsilon}\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}=\tau_k, \tau_k\lambda(\varepsilon)\leqslant t\right)\\ &+\sum_{k=0}^{k_\varepsilon-1}\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\in(\tau_k,\tau_{k+1}), \tau_\pm^{\varepsilon,*}\lambda(\varepsilon)\leqslant t\right)+\mathbb{P}_y\left(\tau_{k_\varepsilon}\leqslant\tau_\pm^{\varepsilon,*}\leqslant t\lambda(\varepsilon)^{-1}\right).\end{aligned} \tag{4.7}$$

The first sum of the previous upper bound contributes the main part to the probability under estimation and the second line creates error terms.
The estimate of the last summand of (4.7) again falls back on the fact that N^{A^T} denotes a Poisson random measure. The error $\mathbb{P}\left(N^{A^T}\left(B_{\varepsilon^{-\rho}}^c(0)\times[0,t\lambda(\varepsilon)^{-1}]\right)\geqslant k_\varepsilon\right)$ vanishes in the small noise limit through choosing $r>\alpha_*(1-\rho)$.
Now concentrate on the main part of (4.7). From the usual factorisation and an application of Lemma 4.22 we can deduce

$$\begin{aligned}&\sum_{k=1}^{k_\varepsilon}\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}=\tau_k,\tau_k\lambda(\varepsilon)\leqslant t\right)\\ &=\sum_{k=1}^{k_\varepsilon}\mathbb{E}_y\,\mathbf{1}_{\{\tau_k\lambda(\varepsilon)\leqslant t\}}\prod_{j=0}^{k-2}\mathbf{1}\left(B^\pm_{\tau_j,\tau_{j+1}}(Y^\varepsilon_{\tau_j})\right)\mathbf{1}\left(C^\pm_{\tau_{k-1},\tau_k}(Y^\varepsilon_{\tau_{k-1}})\right)\\ &\leqslant\sum_{k=1}^{k_\varepsilon}\int_0^t\int_0^{s_k}\cdots\int_0^{s_2}\left[\prod_{j=1}^{k}\left(\sqrt{|\log\varepsilon|}\varepsilon^{\rho\alpha(s_j)-\alpha_*}\nu_{s_j}(B_1^c(0))\right)\mathrm{e}^{-\int_0^{s_k}\sqrt{|\log\varepsilon|}\varepsilon^{\rho\alpha(r)-\alpha_*}\nu_r(B_1^c(0))\mathrm{d}r}\right.\\ &\left.\prod_{j=0}^{k-2}\sup_{y\in\Omega_{\pm,L}(\delta_1^*)}\mathbb{P}(B^\pm_{s_j2T,s_{j+1}2T}(y))\sup_{y\in\Omega_{\pm,L}(\delta_1^*)}\mathbb{P}(C^\pm_{s_{k-1}2T,s_k2T}(y))\right]\mathrm{d}s_1\ldots\mathrm{d}s_k.\end{aligned}$$

A multiplication of the integrand with the factor

$$1=\prod_{j=0}^{k-1}\mathbf{1}_{\{(s_{j+1}-s_j)2T\geqslant R_\varepsilon\gamma\}}+1-\prod_{j=0}^{k-1}\mathbf{1}_{\{(s_{j+1}-s_j)2T\geqslant R_\varepsilon\gamma\}},$$

points the way towards Lemma 4.24 and simultanously produces an error of order $\varepsilon^{\rho\alpha_*-2r}|\log\varepsilon|$ as justified in the proof of the lower bound through estimation of inequality

(4.6). We end up with

$$\begin{aligned}
&\sum_{k=1}^{k_\varepsilon} \mathbb{P}_y\left(\tau_\pm^{\varepsilon,*} = \tau_k, \tau_k\lambda(\varepsilon) \leqslant t\right) \\
&\leqslant \sum_{k=1}^{k_\varepsilon} \int_0^t \int_0^{s_k} \dots \int_0^{s_2} \Bigg[\left(\sqrt{|\log\varepsilon|}\varepsilon^{-\alpha_*}\right)^k \mathrm{e}^{-\int_0^{s_k} \sqrt{|\log\varepsilon|}\varepsilon^{\rho\alpha(r)-\alpha_*}\nu_r(B_1^c(0))\mathrm{d}r} \\
&\prod_{j=1}^{k-1} \left(\varepsilon^{\rho\alpha(s_j)}\nu_{s_j}(B_1^c(0))\,\mathrm{e}^{-\varepsilon^{-p}} + \nu_{s_j}\left(B_{\varepsilon^{-\rho}}^c(0) \cap \frac{1}{\varepsilon}\left(\Omega_{\pm,L_{\delta^*}(L)} - m_\pm\right)\right)\right) \\
&\left(\varepsilon^{\rho\alpha(s_k)}\nu_{s_k}(B_1^c(0))\,\mathrm{e}^{-\varepsilon^{-p}} + \nu_{s_k}\left(\frac{1}{\varepsilon}\left(\Omega_{\pm,L^*}^c(\delta_1^*,\delta_2^*) - m_\pm\right)\right)\right)\Bigg]\mathrm{d}s_1\dots\mathrm{d}s_k + \mathrm{o}_\varepsilon(1)
\end{aligned}$$

with $\mathrm{o}_\varepsilon(1) \to 0$ as $\varepsilon \to 0$. Again extend the finite sum to an infinite series, glance at (ii) of Lemma 3.32, use $\nu_t(\frac{1}{\varepsilon}A) = \varepsilon^{\alpha(t)}\nu_t(A)$ and remember (D1) and (D2) that imply

$$\begin{aligned}
|\nu_t(\Omega_{\pm,L^*}^c(\delta_1^*,\delta_2^*) - m_\pm) - \nu_t(\Omega_{\pm,L}^c(\delta_1^*) - m_\pm)| &\leqslant \frac{\delta^*}{2}, \\
|\nu_t(\Omega_{\pm,L_{\delta^*}(L)} - m_\pm) - \nu_t(\Omega_{\pm,L}^c(\delta_1^*) - m_\pm)| &\leqslant \delta^*.
\end{aligned} \tag{4.8}$$

This proves

$$\begin{aligned}
\sum_{k=1}^{k_\varepsilon} \mathbb{P}_y\left(\tau_\pm^{\varepsilon,*} = \tau_k, \tau_k\lambda(\varepsilon) \leqslant t\right) \leqslant \int_0^t \Bigg[&\sqrt{|\log\varepsilon|}\varepsilon^{\alpha(s)-\alpha_*}\left(\nu_s(\Omega_{\pm,L}^c(\delta_1^*) - m_\pm) + \frac{\delta^*}{2}\right) \\
&\mathrm{e}^{-\int_0^s \sqrt{|\log\varepsilon|}\varepsilon^{\alpha(r)-\alpha_*}\left(\nu_r(\Omega_{\pm,L}^c(\delta_1^*)-m_\pm)-\delta^*\right)\mathrm{d}r}\Bigg]\mathrm{d}s + \mathrm{o}_\varepsilon(1).
\end{aligned} \tag{4.9}$$

Now let us focus on the possibility to leave the basin of attraction between two big jumps and hence consider the second sum of inequality (4.7). Starting within $\Omega_\pm(\delta_1^*,\delta_2^*) \cap O_L^{-\delta_2^*}$ makes an exit from $\Omega_{\pm,L}(\delta_1^*)$ through the small jump process impossible if it lies in the $\frac{1}{2}\varepsilon^{2\gamma}$-neighbourhood of the deterministic process and $\frac{1}{2}\varepsilon^{2\gamma} < \delta_2^*$. This immediately yields

$$\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*} \in (0,\tau_1)\right) \leqslant \mathbb{P}\left(\sup_{t\in[0,\tau_1]} \|\widetilde{Y}_t^\varepsilon(y) - y_t(y)\| \geqslant \frac{1}{2}\varepsilon^{2\beta}\right) = \mathbb{P}\left(E_{0,\tau_1}^c(y)\right) \leqslant \mathrm{e}^{-\varepsilon^{-p}} \tag{4.10}$$

for $\gamma < \frac{\gamma_0}{3}$, $p < p_0$ and $\varepsilon < \varepsilon_0$ with γ_0, p_0 and ε_0 according to Theorem 4.10. For $k \in \mathbb{N}$

$$\begin{aligned}
&\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*} \in (\tau_k,\tau_{k+1}), \tau_\pm^{\varepsilon,*}\lambda(\varepsilon) \leqslant t\right) \\
&\leqslant \mathbb{E}_y \prod_{j=0}^{k-2} \mathbf{1}\left(B_{\tau_j,\tau_{j+1}}^\pm(Y_{\tau_j}^\varepsilon(y))\right)\mathbf{1}\left(\bar{B}_{\tau_{k-1},\tau_k}^\pm(Y_{\tau_{k-1}}^\varepsilon(y))\right)\mathbf{1}\left(F_{\tau_k,\tau_{k+1}}^\pm(Y_{\tau_k}^\varepsilon(y))\right) \\
&\quad + \mathbb{E}_y \prod_{j=0}^{k-2} \mathbf{1}\left(B_{\tau_j,\tau_{j+1}}^\pm(Y_{\tau_j}^\varepsilon(y))\right)\mathbf{1}\left(\hat{B}_{\tau_{k-1},\tau_k}^\pm(Y_{\tau_{k-1}}^\varepsilon(y))\right)\mathbf{1}_{\{\tau_k\lambda(\varepsilon)\leqslant t\}}
\end{aligned} \tag{4.11}$$

is true. Estimate the first summand through arguments presented above

$$\begin{aligned}
&\mathbb{E}_y \prod_{j=0}^{k-2} \mathbf{1}\left(B^{\pm}_{\tau_j,\tau_{j+1}}(Y^{\varepsilon}_{\tau_j}(y))\right) \mathbf{1}\left(\bar{B}^{\pm}_{\tau_{k-1},\tau_k}(Y^{\varepsilon}_{\tau_{k-1}}(y))\right) \mathbf{1}\left(F^{\pm}_{\tau_k,\tau_{k+1}}(Y^{\varepsilon}_{\tau_k}(y))\right) \\
&\leqslant \mathbb{E}_y \left(\prod_{j=0}^{k-2} \mathbf{1}\left(B^{\pm}_{\tau_j,\tau_{j+1}}(Y^{\varepsilon}_{\tau_j}(y))\right) \mathbf{1}\left(\bar{B}^{\pm}_{\tau_{k-1},\tau_k}(Y^{\varepsilon}_{\tau_{k-1}}(y))\right) \mathbb{E}\left(E^{c}_{\tau_k,\tau_{k+1}}(Y^{\varepsilon}_{\tau_k}(y)) \Big| \mathcal{F}_{\tau_k}\right)\right).
\end{aligned}$$

Due to the strong Markov property and Theorem 4.10 the conditional expectation above is exponentially small in ε. For the last summand in formula (4.11) revive the arguments of the possibility to exit the well at τ_k. Apply Lemma 4.22, introduce the factor demanding inter-jump times bigger than the logarithmic return time, control the error terms and use Lemma 4.24 (iii) and (v) to end up with

$$\begin{aligned}
&\sum_{k=1}^{k_\varepsilon-1} \mathbb{E}_y \prod_{j=0}^{k-2} \mathbf{1}\left(B^{\pm}_{\tau_j,\tau_{j+1}}(Y^{\varepsilon}_{\tau_j}(y))\right) \mathbf{1}\left(\hat{B}^{\pm}_{\tau_{k-1},\tau_k}(Y^{\varepsilon}_{\tau_{k-1}}(y))\right) \mathbf{1}_{\{\tau_k\lambda(\varepsilon)\leqslant t\}} \\
&\leqslant \int_0^t \Bigg[\sqrt{|\log\varepsilon|}\varepsilon^{\alpha(s)-\alpha_*}\left(\nu_s(\Gamma_{L^*}(\delta_1^*,\delta_2^*,\delta_2^*)-m_\pm)+\nu_s(O^c_{L^*}-m_\pm)\right) \\
&\qquad \mathrm{e}^{-\int_0^s \sqrt{|\log\varepsilon|}\varepsilon^{\alpha(r)-\alpha_*}\nu_r(\Omega^c_{\pm,L_{\delta_1^*}(L)}-m_\pm)\mathrm{d}r}\Bigg]\mathrm{d}s + \mathrm{o}_\varepsilon(1).
\end{aligned}$$

From the estimate above combined with the careful choice of δ_1^*,δ_2^* and L_* and all previous estimates we can derive

$$\begin{aligned}
&\mathbb{P}_y\left(\tau^{\varepsilon,*}_{\pm}\lambda(\varepsilon)\leqslant t\right) \\
&\leqslant \mathrm{o}_\varepsilon(1) + \int_0^t \sqrt{|\log\varepsilon|}\varepsilon^{\alpha(s)-\alpha_*}(\nu_s(\Omega^c_{\pm,L}-m_\pm)+\delta^*)\,\mathrm{e}^{-\int_0^s\sqrt{|\log\varepsilon|}\varepsilon^{\alpha(r)-\alpha_*}(\nu_r(\Omega^c_{\pm,L}-m_\pm)-\delta^*)\mathrm{d}r}\mathrm{d}s \\
&\leqslant \delta^* + \left(1+\frac{2\delta^*}{\min_{s\in[0,1]}\nu_s(\Omega^c_{\pm,L}-m_\pm)-\delta^*}\right)\left(1-\mathrm{e}^{-\int_0^t\sqrt{|\log\varepsilon|}\varepsilon^{\alpha(s)-\alpha_*}\nu_s(\Omega^c_{\pm,L}-m_\pm)\mathrm{d}s}\right) \\
&\leqslant \delta^*\left(1+\frac{2}{\min_{s\in[0,1]}\nu_s(\Omega^c_{\pm,L}-m_\pm)-\delta^*}\right) + \left(1-\mathrm{e}^{-\int_0^t\sqrt{|\log\varepsilon|}\varepsilon^{\alpha(s)-\alpha_*}\nu_s(\Omega^c_{\pm,L}-m_\pm)\mathrm{d}s}\right)
\end{aligned}$$

for sufficiently small ε and $0<\delta^*\leqslant 1\wedge\frac{1}{2}\min_{s\in[0,1]}\nu_s(\Omega^c_{\pm,L}-m_\pm)$. □

Proof of the Theorem 4.18:

Proof. According to Lemma 4.20 it remains to prove

$$\left|\exp\left(-\int_0^t \varepsilon^{\alpha(r)-\alpha_*}\sqrt{|\log\varepsilon|}\nu_r(\Omega^c_{\pm,L}-m_\pm)\,\mathrm{d}r\right)-\mathrm{e}^{-g(t)c_\pm}\right|\leqslant\delta^*$$

with g seen in Figure 4.2. Due to (D1) we can replace $\nu_r(\Omega^c_{\pm,L}-m_\pm)$ by $\nu_r(\Omega^c_{\pm}-m_\pm)$ which produces errors of order δ^*. It is left to prove

$$\lim_{\varepsilon\to0}\frac{\sqrt{|\log\varepsilon|}}{c_\pm}\int_0^t\varepsilon^{\alpha(r)-\alpha_*}\nu_r(\Omega^c_\pm-m_\pm)\,\mathrm{d}r = g(t)$$

for $t \geqslant 0$. Due to the 1-periodicity of the integrand and g it suffices to concentrate on $t \in [0,1]$. Define $v_r = \nu_r(\Omega_\pm^c - m_\pm)$. Because of Lemma 2.47 we obtain the limit

$$\lim_{\varepsilon\to 0} \frac{\sqrt{|\log\varepsilon|}}{c_\pm} \int_0^1 \varepsilon^{\alpha(r)-\alpha_*} v_r \,\mathrm{d}r = 1 = g(1)$$

and analogously we justify

$$\lim_{\varepsilon\to 0} \frac{\sqrt{|\log\varepsilon|}}{c_\pm} \int_0^a \varepsilon^{\alpha(r)-\alpha_*} v_r \,\mathrm{d}r = \frac{1}{2} = g(a).$$

If $t < a$ define $b = \min_{r\in[0,t]} \alpha(r) > \alpha_*$ and we get

$$\lim_{\varepsilon\to 0} \frac{\sqrt{|\log\varepsilon|}}{c_\pm} \int_0^t \varepsilon^{\alpha(r)-\alpha_*} v_r \,\mathrm{d}r \leqslant \lim_{\varepsilon\to 0} \frac{\sqrt{|\log\varepsilon|}}{c_\pm} t\, \varepsilon^{b-\alpha_*} \max_{r\in[0,t]} v_r = 0 = g(t).$$

If $t \in (a,1)$ we use the methods of the proof of Lemma 2.47. Assume $\alpha(r) - \alpha_* \leqslant (\alpha''(a) + \delta)(r-a)^2\frac{1}{2}$ for $|r-a| < \xi$ and $v_r \geqslant v_a - \delta$ for $|r-a| < \tilde{\xi}$ and for $\bar{\xi} = \xi \wedge \tilde{\xi} \wedge (t-a)$ estimate

$$\int_0^t \varepsilon^{\alpha(r)-\alpha_*} v_r \,\mathrm{d}r \geqslant \int_{a-\bar{\xi}}^{a+\bar{\xi}} \mathrm{e}^{-|\log\varepsilon|(\alpha(r)-\alpha_*)} v_r \,\mathrm{d}r \geqslant (v_a - \delta) \int_{a-\bar{\xi}}^{a+\bar{\xi}} \mathrm{e}^{-(\alpha''(a)+\delta)|\log\varepsilon|(r-a)^2/2} \mathrm{d}r.$$

Now substitute $u = \sqrt{(\alpha''(a)+\delta)|\log\varepsilon|}(r-a)$. This yields the lower bound

$$\begin{aligned}
\lim_{\varepsilon\to 0} \frac{\sqrt{|\log\varepsilon|}}{c_\pm} \int_0^t \varepsilon^{\alpha(r)-\alpha_*} v_r \,\mathrm{d}r &\geqslant \lim_{\varepsilon\to 0} \frac{\sqrt{|\log\varepsilon|}}{c_\pm} (v_a-\delta) \int_{a-\bar{\xi}}^{a+\bar{\xi}} \mathrm{e}^{-(\alpha''(a)+\delta)|\log\varepsilon|(r-a)^2/2} \mathrm{d}r \\
&\geqslant \lim_{\varepsilon\to 0} \frac{\sqrt{\alpha''(a)}(v_a-\delta)}{\sqrt{(\alpha''(a)+\delta)}v_a} \int_{-\bar{\xi}\sqrt{(\alpha''(a)+\delta)|\log\varepsilon|}}^{\bar{\xi}\sqrt{(\alpha''(a)+\delta)|\log\varepsilon|}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-u^2/2} \mathrm{d}u \\
&= \frac{\sqrt{\alpha''(a)}(v_a-\delta)}{\sqrt{(\alpha''(a)+\delta)}v_a}.
\end{aligned}$$

An upper bound is guaranteed as follows. For arbitrary $\delta < \delta_0$ with $\alpha''(a) - \delta_0 > 0$ there are $\xi, \tilde{\xi} > 0$ such that $\alpha(r) - \alpha_* \geqslant (\alpha''(a) - \delta)(r-a)^2\frac{1}{2}$ for $|r-a| \leqslant \xi$ and $v_r \leqslant v_a + \delta$ for $|r-a| \leqslant \tilde{\xi}$. Choose $\bar{\xi} = \xi \wedge \tilde{\xi} \wedge (t-a)$ and define $b = \min_{r\in[a+\bar{\xi},t]} \alpha(r) > \alpha_*$ then we prove the upper bound

$$\begin{aligned}
\int_0^t \varepsilon^{\alpha(r)-\alpha_*} v_r \,\mathrm{d}r &= \int_{a-\bar{\xi}}^{a+\bar{\xi}} \varepsilon^{\alpha(r)-\alpha_*} v_r \,\mathrm{d}r + \int_{a+\bar{\xi}}^{t} \varepsilon^{\alpha(r)-\alpha_*} v_r \,\mathrm{d}r \\
&\leqslant (v_a+\delta) \int_{a-\bar{\xi}}^{a+\bar{\xi}} \mathrm{e}^{-(\alpha''(a)-\delta)|\log\varepsilon|(r-a)^2/2} \mathrm{d}r + (t-a)\varepsilon^{b-\alpha_*} \max_{s\in[0,1]} v_s.
\end{aligned}$$

Since $\sqrt{|\log\varepsilon|}\varepsilon^{b-\alpha_*}$ is a null sequence as $\varepsilon \to 0$ we obtain

$$\begin{aligned}
\lim_{\varepsilon\to 0} \frac{\sqrt{|\log\varepsilon|}}{c_\pm} \int_0^t \varepsilon^{\alpha(r)-\alpha_*} v_r \,\mathrm{d}r &\leqslant \lim_{\varepsilon\to 0} \frac{\sqrt{\alpha''(a)}(v_a+\delta)}{\sqrt{(\alpha''(a)-\delta)}v_a} \int_{-\bar{\xi}\sqrt{(\alpha''(a)-\delta)|\log\varepsilon|}}^{\bar{\xi}\sqrt{(\alpha''(a)-\delta)|\log\varepsilon|}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-u^2/2} \mathrm{d}u \\
&= \frac{\sqrt{\alpha''(a)}(v_a+\delta)}{\sqrt{(\alpha''(a)-\delta)}v_a}.
\end{aligned}$$

Combining the upper and lower bound proves the limit $1 = g(t)$ for $t \in (a,1)$. $\square$

4.4.2 Exit from a domain of attraction

In this subsection we will get rid of the localization parameters δ^*, δ_1^*, δ_2^* and L^* and consider the exit time of a whole basin of attraction.

Definition 4.25. *Define the exit time of Y^ε from the domain of attraction $\Omega_\pm$ through*

$$\tau_\pm^\varepsilon = \inf\left\{t \geqslant 0 : Y_t^\varepsilon(y) \notin \Omega_\pm\right\}.$$

The theory about Lévy-driven jump diffusions suggests that the mean exit from a well occurs at a time polynomially big in ε. Different polynomial times are examined in the following assertion.

Lemma 4.26. *Assume the initial value y of the jump diffusion Y^ε is chosen from $\Omega_\pm$.*

(i) If $\mu \leqslant \alpha_*$ for all $t \geqslant 0$ we have $\lim_{\varepsilon\to 0} \mathbb{P}_y(\tau_\pm^\varepsilon \varepsilon^\mu > t) = 1$.

(ii) For $\mu > \alpha^*$ and all $t \geqslant 0$ it holds $\lim_{\varepsilon\to 0} \mathbb{P}_y(\tau_\pm^\varepsilon \varepsilon^\mu \leqslant t) = 1$.

(iii) Assume $\mu \in (\alpha_*, \alpha^*]$ and define $u = \inf\{s \geqslant 0 : \alpha(s) = \mu\}$. Then the following holds true

- *If $\alpha(0) > \mu$ two cases must be distinguished:*
  - *a) for $t < u$: $\lim_{\varepsilon\to 0} \mathbb{P}_y(\tau_\pm^\varepsilon \varepsilon^\mu > t) = 1$,*
  - *b) for $t > u$: $\lim_{\varepsilon\to 0} \mathbb{P}_y(\tau_\pm^\varepsilon \varepsilon^\mu \leqslant t) = 1$.*
- *If $\alpha(0) < \mu$, then $\lim_{\varepsilon\to 0} \mathbb{P}_y(\tau_\pm^\varepsilon \varepsilon^\mu \leqslant t) = 1$.*

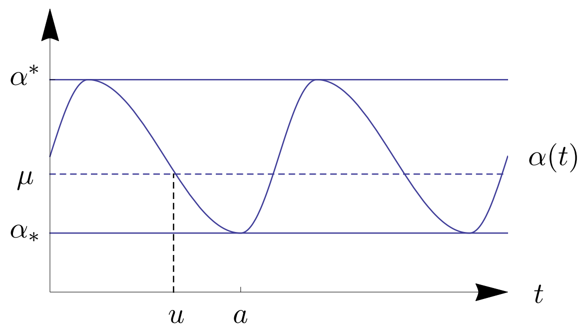


Figure 4.3: On example for the case (iii) of Lemma 4.26 with $\mu < \alpha(0)$.

Proof. An upper bound for the probability $\mathbb{P}_y(\tau_\pm^\varepsilon \varepsilon^\mu \leqslant t)$ falls back on the simple fact that $\tau_\pm^{\varepsilon,*} \leqslant \tau_\pm^\varepsilon$ since an exit of a bounded reduced domain always occurs before an exit of the whole domain of attraction. That is why $\mathbb{P}_y(\tau_\pm^\varepsilon \varepsilon^\mu \leqslant t) \leqslant \mathbb{P}_y(\tau_\pm^{\varepsilon,*} \varepsilon^\mu \leqslant t)$ is true. Analogously to Lemma 4.20 the estimate

$$\left|\mathbb{P}_y(\tau_\pm^{\varepsilon,*}\varepsilon^\mu \leqslant t) - \left(1 - \exp\left(-\int_0^t \varepsilon^{\alpha(r)-\mu}\nu_r(\Omega_{\pm,L}^c - m_\pm)\,\mathrm{d}r\right)\right)\right| \leqslant \delta^* \qquad (4.12)$$

is valid for small ε. The involved integral is of order $\varepsilon^{\alpha_*-\mu}|\log\varepsilon|^{-1/2}$ (Lemma 2.47) and thus converges to zero for $\mu \leqslant \alpha_*$ respectively to infinity if $\mu > \alpha^*$. For $\mu \leqslant \alpha_*$ the probability $\mathbb{P}_y(\tau_\pm^{\varepsilon,*}\varepsilon^\mu \leqslant t)$ is smaller than $c_1\delta^*$ for some $c_1 > 0$ and small ε. This proves $\lim_{\varepsilon\to 0}\mathbb{P}_y(\tau_\pm^\varepsilon\varepsilon^\mu \leqslant t) = 0$.
If $\mu > \alpha^*$ intersecting of $\left\{\tau_\pm^\varepsilon\varepsilon^\mu \leqslant t\right\}$ with $\left\{\tau_\pm^\varepsilon = \tau_\pm^{\varepsilon,*}\right\}$ yields

$$\mathbb{P}_y(\tau_\pm^\varepsilon\varepsilon^\mu \leqslant t) \geqslant \mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\varepsilon^\mu \leqslant t, Y^\varepsilon_{\tau_\pm^{\varepsilon,*}} \notin \Omega_\pm\right). \tag{4.13}$$

We will repeat the steps of the proof of Lemma 4.20 (lower bound) and for this redefine the set $C^\pm_{r,u}(y)$ given in Definition 4.21 as

$$C^\pm_{r,u}(y) = \left\{\widetilde{Y}^\varepsilon_{r,v}(y) \in \Omega_{\pm,L}(\delta_1^*), v \in [0, u-r), \widetilde{Y}^\varepsilon_{r,u-r}(y) + W_u^\varepsilon \notin \Omega_\pm\right\}$$

for arbitrary $0 \leqslant r \leqslant u$ and W_u^ε with law $\nu_{u/2T}\left(B^c_{\varepsilon^{-\rho}}(0) \cap \frac{1}{\varepsilon}\cdot\right)/\nu_{u/2T}(B^c_{\varepsilon^{-\rho}}(0))$. Because of $\delta_1^* > \delta_2^*$ the set $(\Omega_\pm^c\backslash\Gamma(\delta_1^*))^{+\delta_2^*}$ lies within $\Omega_\pm^c$. This justifies an estimate of a similar type as those in Lemma 4.24 for the set $C^\pm_{s2T,(s+t)2T}(y)$ with $2T = \varepsilon^{-\mu}$ given above:

$$\begin{aligned}
&\inf_{y\in\Omega_\pm(\delta_1^*,\delta_2^*)\cap O_L^{-\delta_2^*}} \mathbb{P}(C^\pm_{s2T,(s+t)2T}(y))\mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}\\
&\geqslant \mathbf{1}_{\{t2T\geqslant R_\varepsilon\gamma\}}\left(1 - \mathrm{e}^{-\varepsilon^{-p}}\right)\frac{\nu_{s+t}\left(\frac{1}{\varepsilon}\left(\Omega_\pm^c\backslash\Gamma(\delta_1^*) - m_\pm\right)\right)}{\varepsilon^{\rho\alpha(s+t)}\nu_{s+t}(B_1^c(0))}.
\end{aligned}$$

All preliminaries that are necessary for an estimate as in Lemma 4.20 are stated. From this method we can derive

$$\begin{aligned}
\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\varepsilon^\mu \leqslant t, Y^\varepsilon_{\tau_\pm^{\varepsilon,*}} \notin \Omega_\pm\right) \geqslant \mathrm{o}_\varepsilon(1) + \int_0^t \Bigg[&\varepsilon^{\alpha(s)-\mu}\nu_s(\Omega_\pm^c\backslash\Gamma(\delta_1^*) - m_\pm)\\
&\exp\left(-\int_0^s \varepsilon^{\alpha(r)-\mu}\nu_r(\Omega^c_{\pm,L}(\delta_1^*,\delta_2^*,\delta_2^*) - m_\pm)\,\mathrm{d}r\right)\Bigg]\mathrm{d}s,
\end{aligned}$$

where $\mathrm{o}_\varepsilon(1)$ denotes an error term that tends to zero as ε does so. Replacing the complements of all bounded and reduced domains by $\Omega_\pm^c$ produces errors of order δ^* due to (D1) and (D2) and we get the lower bound

$$\mathbb{P}_y\left(\tau_\pm^{\varepsilon,*}\varepsilon^\mu \leqslant t, Y^\varepsilon_{\tau_\pm^{\varepsilon,*}} \notin \Omega_\pm\right) \geqslant -c_2\delta^* + 1 - \exp\left(-\int_0^t \varepsilon^{\alpha(r)-\mu}\nu_r(\Omega_\pm^c - m_\pm)\,\mathrm{d}r\right) \tag{4.14}$$

for some $c_2 > 0$. Since $\mu > \alpha^*$ and inequality (4.13) holds assertion (ii) follows.
Assume $\mu \in (\alpha_*, \alpha^*]$, $\alpha(0) > \mu$ and $t < u$. Due to the validity of inequality (4.12) and $\mu < \min_{r\in[0,t]}\alpha(r)$ we are able to estimate for some $c_3 > 0$

$$\begin{aligned}
\mathbb{P}_y(\tau_\pm^{\varepsilon,*}\varepsilon^\mu \leqslant t) &\leqslant \delta^* + \left(1 - \exp\left(-t\varepsilon^{\min_{r\in[0,t]}\alpha(r)-\mu}\left(\max_{s\in[0,1]}\nu_s(\Omega^c_{\pm,L} - m_\pm) + \delta^*\right)\right)\right)\\
&\leqslant c_3\delta^*
\end{aligned}$$

for small ε. Because $\tau_{\pm}^{\varepsilon,*} \leqslant \tau_{\pm}^{\varepsilon}$ is true assertion a) is proven. If $t > u$, define $v = \inf\{s > u : \alpha(s) = \mu\} \wedge t$ and use inequalities (4.13) and (4.14) to estimate

$$\begin{aligned}\mathbb{P}_y(\tau_{\pm}^{\varepsilon}\varepsilon^{\mu} \leqslant t) &\geqslant -c_2\delta^* + 1 - \exp\Big(-\int_{u+\delta^*}^{v-\delta^*} \varepsilon^{\alpha(v-\delta^*)\vee\alpha(u+\delta^*)-\mu} \min_{s\in[0,1]} \nu_s(\Omega_{\pm}^c - m_{\pm})\,\mathrm{d}x\Big)\\ &\geqslant 1 - c_4\delta^*\end{aligned}$$

for some $c_4 > 0$ and small ε, because $\mu > \alpha(v-\delta^*) \vee \alpha(u+\delta^*)$. The case $\alpha(0) < \mu$ can be treated analogously. We prove

$$\begin{aligned}&\mathbb{P}_y(\tau_{\pm}^{\varepsilon}\varepsilon^{\mu} \leqslant t)\\ &\geqslant -c_2\delta^* + 1 - \exp\left(-\int_0^{t\wedge(u-\delta)} \varepsilon^{\max_{r\in[0,t\wedge(u-\delta)]}\alpha(r)-\mu} \min_{s\in[0,1]} \nu_s(\Omega_{\pm}^c - m_{\pm})\,\mathrm{d}x\right).\end{aligned}$$

Since $\mu > \max_{r\in[0,t\wedge(u-\delta)]}\alpha(r)$ the lower bound above is bigger than $1 - c_5\delta^*$ for some $c_5 > 0$. This finishes the proof of (iii). $\square$

Definition 4.27. *Let $Z = (Z_n)_{n\in\mathbb{N}}$ be a discrete time Markov chain with values in $\{-1, 1\}$, initial value $Z_0 = z_0 \in \{-1, 1\}$ and transition matrix*

$$P = \begin{pmatrix} \mathrm{e}^{-c_-} & 1-\mathrm{e}^{-c_-}\\ 1-\mathrm{e}^{-c_+} & \mathrm{e}^{-c_+}\end{pmatrix}.$$

Its time-continuous extension denoted by Z^c is given by

$$Z_t^c = \begin{cases} Z_0, & t < a,\\ Z_k, & t \in [a+k-1, a+k),\ k \in \mathbb{N}\end{cases}$$

and its jump times are $(\tau_n^{Z,c})_{n\in\mathbb{N}}$.

Theorem 4.28. *Assume $y \in \Omega_{\pm}$. In the small noise limit the exit time $\tau_{\pm}^{\varepsilon}$ of Y^{ε} is distributed as follows*

$$\lim_{\varepsilon\to 0} \mathbb{P}_y(\tau_{\pm}^{\varepsilon}\lambda(\varepsilon) \leqslant t) = 1 - \mathrm{e}^{-c_{\pm}g(t)}$$

for $t \geqslant 0$. Let $z_0(y) = -\mathbf{1}_{\Omega_-}(y) + \mathbf{1}_{\Omega_+}(y)$. Then for all $\delta > 0$ and $k \in \mathbb{N}$ we get

(i)

$$\begin{aligned}&\lim_{\varepsilon\to 0} \mathbb{P}_y(\tau_{\pm}^{\varepsilon}\lambda(\varepsilon) \in (a+k-1-\delta, a+k-1+\delta])\\ &= \mathrm{e}^{-(k-1)c_{\pm}}(1-\mathrm{e}^{-c_{\pm}})\\ &= \mathbb{P}_{z_0(y)}(\tau_1^{Z,c} = a+k-1),\end{aligned}$$

(ii)

$$\lim_{\varepsilon\to 0} \mathbb{P}_y(\tau_{\pm}^{\varepsilon}\lambda(\varepsilon) \in (a+k-1+\delta, a+k-\delta]) = 0 = \mathbb{P}_{z_0(y)}(\tau_1^{Z,c} \in (a+k-1, a+k)).$$

Remark 4.29. Theorem 4.28 states that an exit of $Y^\varepsilon_{t/\lambda(\varepsilon)}$ around $a+k-1$, $k\in\mathbb{N}$, which is a minimum position of α, in the small noise limit is as probable as a first jump of Z after k time units. In comparison to that an exit of the scaled jump diffusion between two minimum positions of α is extremely small. This coincides with the jump probability of the continuous-time extension of Z to jump between $a+k-1$ and $a+k$.

Proof. Due to the definition of a limit we assume $\delta^*>0$ is a small and fixed level of deviation that need not to be overcome by the deviation of the probability under estimation and its limit if ε is chosen appropriately small. Recall the inequality $\tau^{\varepsilon,*}_\pm \leqslant \tau^\varepsilon_\pm$. Additionally due to Theorem 4.18 there exist $\varepsilon_0(\delta^*)>0$ and $C^*>0$ with

$$\mathbb{P}_y(\tau^\varepsilon_\pm\lambda(\varepsilon)\leqslant t)\leqslant \mathbb{P}_y(\tau^{\varepsilon,*}_\pm\lambda(\varepsilon)\leqslant t)\leqslant 1-\mathrm{e}^{-g(t)c_\pm}+C^*\delta^*.$$

We continue searching for a corresponding lower bound. Intersecting of $\left\{\tau^\varepsilon_\pm\lambda(\varepsilon)\leqslant t\right\}$ with $\left\{\tau^\varepsilon_\pm=\tau^{\varepsilon,*}_\pm\right\}$ yields

$$\mathbb{P}_y(\tau^\varepsilon_\pm\lambda(\varepsilon)\leqslant t)\geqslant \mathbb{P}_y\left(\tau^{\varepsilon,*}_\pm\lambda(\varepsilon)\leqslant t, Y^\varepsilon_{\tau^{\varepsilon,*}_\pm}\notin\Omega_\pm\right).$$

We will repeat the methods of the proof of (ii) of Lemma 4.26 with $2T=\lambda(\varepsilon)^{-1}$ to get

$$\mathbb{P}_y\left(\tau^{\varepsilon,*}_\pm\lambda(\varepsilon)\leqslant t, Y^\varepsilon_{\tau^{\varepsilon,*}_\pm}\notin\Omega_\pm\right)\geqslant -C\delta^*+1-\exp\left(-\int_0^t \varepsilon^{\alpha(r)-\alpha_*}\sqrt{|\log\varepsilon|}\nu_r(\Omega^c_\pm-m_\pm)\,\mathrm{d}r\right)$$

for some $C>0$. Again use Laplace's method (Lemma 2.47) to prove the small noise limit $-c_\pm g(t)$ of the exponent.
This yields assertion (i) because

$$\begin{aligned}\lim_{\varepsilon\to 0}\mathbb{P}_y(\tau^\varepsilon_\pm\lambda(\varepsilon)\in(a+k-1-\delta,a+k-1+\delta]) &= \mathrm{e}^{-c_\pm g(a+k-1-\delta)}-\mathrm{e}^{-c_\pm g(a+k-1+\delta)}\\ &=\mathrm{e}^{-c_\pm(k-1)}(1-\mathrm{e}^{-c_\pm})\\ &=\mathbb{P}_{z_0(y)}(Z_1=\ldots=Z_{k-1}\neq Z_k)\end{aligned}$$

and (ii) can be justified through $g(a+k-\delta)=g(a+k-1+\delta)$ and the impossibility of Z^c to jump between $a+k-1$ and $a+k$. $\square$

Bibliography

[1] H. Amann. *Gewöhnliche Differentialgleichungen.* De-Gruyter-Lehrbuch. De Gruyter, Berlin, 1995.

[2] D. Applebaum. *Lévy Processes and Stochastic Calculus.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2009.

[3] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani. A theory of stochastic resonance in climatic change. *SIAM Journal on Applied Mathematics*, 43(3):565–578, 1983.

[4] N. Berglund. *Adiabatic Dynamical Systems and Hysteresis.* Phd thesis, École polytechnique fédérale de Lausanne, 1998.

[5] N. Berglund and B. Gentz. Universality of first-passage and residence-time distributions in non-adiabatic stochastic resonance. *Europhysics Letters*, 70:1–7, 2005.

[6] N. Berglund and B. Gentz. *Noise-Induced Phenomena in Slow-Fast Dynamical Systems: A Sample-Paths Approach.* Probability and Its Applications. Springer, London, 2006.

[7] N. Bingham, C. Goldie, and J. Teugels. *Regular Variation*, volume 27 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1989.

[8] B. Böttcher. Feller evolution systems: generators and approximation. *Stochastics and Dynamics*, 14(3):1350025, 2014.

[9] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)*, 6(4):399–424, 2004.

[10] D. Brockmann and I. Sokolov. Lévy flights in external force fields: from models to equations. *Chemical Physics*, 284(1-2):409–421, 2002.

[11] M. I. Budyko. The effect of solar radiation variations on the climate of the Earth. *Tellus*, 21(5):611–619, 1969.

[12] A. Cartea and S. Howison. Option Pricing with Lévy-Stable Processes Generated by Lévy-Stable Integrated Variance. *Quantitative Finance*, 9(4):397–409, 2009.

[13] P. D. Ditlevsen. Observation of α-stable noise induced millennial climate changes from an ice-core record. *Geophysical Research Letters*, 26(10):1441–1444, 1999.

[14] S. Ethier and T. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley Series in Probability and Statistics. John Wiley & Sons, Hoboken, 2009.

[15] M. I. Freidlin. Quasi-deterministic approximation, metastability and stochastic resonance. *Physica D: Nonlinear Phenomena*, 137(34):333–352, 2000.

[16] M. I. Freidlin and A. Wentzell. *Random Perturbations of Dynamical Systems*. Grundlehren der mathematischen Wissenschaften. Springer, New York, 1998.

[17] P. G. Georgiou, S. Member, P. Tsakalides, and C. Kyriakakis. Alpha-Stable Modeling of Noise and Robust Time-Delay Estimation in the Presence of Impulsive Noise. *IEEE Transactions on Multimedia*, 1(3):291–301, 1999.

[18] J. Hale. *Ordinary Differential Equations*, volume 21 of *Dover Books on Mathematics Series*. Dover Publications, Mineola, New York, 2009.

[19] S. Herrmann and P. Imkeller. Barrier crossings characterize stochastic resonance. *Stochastics and Dynamics*, 2(3):413–436, 2002.

[20] S. Herrmann, P. Imkeller, I. Pavlyukevich, and D. Peithmann. *Stochastic Resonance: A Mathematical Approach in the Small Noise Limit*, volume 194 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, Rhode Island, 2013.

[21] H. Hult and F. Lindskog. On Regular Variation for Infinitely Divisible Random Vectors and Additive Processes. *Advances in Applied Probability*, 38(1):134–148, 2006.

[22] H. Hult and F. Lindskog. Regular Variation for Measures on Metric Spaces. *Publications de l'Institut Mathématique (Beograd), Nouvelle Série*, 80(94)(100):121–140, 2006.

[23] P. Imkeller. Energy Balance Models – viewed from Stochastic Dynamics. In P. Imkeller and J.-S. von Storch, editors, *Stochastic climate models. Proceedings of a workshop, Chorin, Germany, Summer 1999*, volume 49 of Progress in Probability, pages 213–240, Basel, 2001. Birkhäuser.

[24] P. Imkeller and I. Pavlyukevich. Stochastic Resonance in Two-state Markov Chains. *Arch. Math. (Basel)*, 77(1):107–115, 2001.

[25] P. Imkeller and I. Pavlyukevich. First exit times of SDEs driven by stable Lévy processes. *Stochastic Processes and their Applications*, 116(4):611 – 642, 2006.

[26] P. Imkeller and I. Pavlyukevich. Metastable behaviour of small noise Lévy-driven diffusions. *ESAIM: Probability and Statistics*, 12:412–437, 2008.

[27] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*, volume 288 of *Grundlehren der mathematischen Wissenschaften.* Springer, Berlin, second edition, 2010.

[28] S. J. Johnson. High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature*, 431(7005):147–151, 2004.

[29] A. Kabašinskas, S. T. Rachev, L. Sakalauskas, W. Sun, and I. Belovas. Alpha-stable paradigm in financial markets. *Journal of Computational Analysis & Applications*, 11(4):641–668, 2009.

[30] O. Kallenberg. *Foundations of Modern Probability.* Probability and Its Applications. Springer, New York, second edition, 2010.

[31] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.

[32] T. G. Kurtz, É. Pardoux, and P. Protter. Stratonovich stochastic differential equations driven by general semimartingales. *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, 31(2):351–377, 1995.

[33] F. Lindskog. *Multivariate Extremes and Regular Variation for Stochastic Processes.* Phd thesis, Eidgenössische Technische Hochschule Zürich, 2004.

[34] X. Mao. *Stochastic Differential Equations and Applications.* Woodhead Publishing, Philadelphia, 2007.

[35] C. Nicolis. Stochastic aspects of climatic transitions response to a periodic forcing. *Tellus*, 34(1):1–9, 1982.

[36] F. W. J. Olver. *Asymptotics and special functions.* Computer science and applied mathematics. CRC Press, Boca Raton, Florida, 2010.

[37] I. Pavlyukevich. *Stochastic Resonance.* Phd thesis, Humboldt-Universität zu Berlin, 2002.

[38] I. Pavlyukevich. First exit times of solutions of stochastic differential equations driven by multiplicative Lévy noise with heavy tails. *Stochastics and Dynamics*, 11(2-3):495–519, 2011.

[39] P. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Applications of Mathematics.* Springer, Berlin, second edition, 2004.

[40] S. Resnick. On the foundation of multivariate heavy-tail analysis. *Journal of Applied Probability*, 41A:191–212, 2004.

[41] G. Samorodnitsky and M. Grigoriu. Tails of solutions of certain nonlinear stochastic differential equations driven by heavy tailed Lévy motions. *Stochastic Processes and their Applications*, 105(1):69–97, 2003.

[42] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[43] W. Sellers. A Global Climatic Model Based on the Energy Balance of the Earth-Atmosphere System. *Journal of applied meteorology*, 8:392–400, 1969.

[44] P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC Press, London, 2004.

[45] M. Wilke and J. Prüss. *Gewöhnliche Differentialgleichungen und dynamische Systeme*. Grundstudium Mathematik. Springer, Basel, 2010.

[46] R. Wong. *Asymptotic Approximations of Integrals*, volume 34 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, 2001.

List of notations

Notation related to the ordinary differential equations:

$U(x,t)$	time-periodic double-well potential (page 12)
$m_\pm(t)$	minimum positions of U
$p_\pm(t)$	periodic solutions (page 13)
$V(x)$	double-well potential (page 11)
$m_\pm$	minimum positions of V

Notation related to the perturbation terms:

ε	noise amplitude
$(L_t)_{t\geqslant 0}$	Lévy process (pp. 14, 20)
$(\xi_t)_{t\geqslant 0}$	small jump part of L (page 52)
$(\eta_t)_{t\geqslant 0}$	big jump part of L (page 52)
g	matrix-valued function (page 20)
ν	Lévy measure of L (pp. 14, 20)
$\diamond$	canonical Marcus integral (page 21)
$(A_t^T)_{t\geqslant 0}$	additive process (pp. 14, 22)
$(\widetilde{A}_t^T)_{t\geqslant 0}$	small jump part of A^T (page 22)
N^{A^T}	Poisson random measure associated with A^T (pp. 17, 22)
$\widetilde{N}^{A^T}$	compensated Poisson random measure (page 23)
ν_t	Lévy measures associated with A^T (pp. 14, 22)
$\alpha(t)$	periodic function (page 23)
α_*	minimal value of the function α (page 23)
R_{ε^γ}	logarithmic return time (pp. 52, 82)

Two-state Markov chains approximating the diffusions:

$(K_t^\varepsilon)_{t\geqslant 0}$	Chain with piecewise constant Q-matrix approximating $(X_t^\varepsilon)_{t\geqslant 0}$ (page 37)
$(C_t^\varepsilon)_{t\geqslant 0}$	Chain with continuous Q-matrix approximating $(X_t^\varepsilon)_{t\geqslant 0}$ (page 45)
$(\mathcal{C}_t^\varepsilon)_{t\geqslant 0}$	Chain with continuous Q-matrix approximating $(Y_t^\varepsilon)_{t\geqslant 0}$ (page 76)

Jump diffusions :

$(X_t^\varepsilon)_{t\geqslant 0}$	jump diffusion with time-periodic drift perturbed by multiplicative Lévy noise using the Itô integral (page 36)
$(Z_t^\varepsilon)_{t\geqslant 0}$	jump diffusion with time-periodic drift perturbed by multiplicative Lévy noise using the canonical Marcus integral (page 36)
$(X_{s,t}^\xi)_{t\geqslant 0}$	small jump process belonging to $(X_t^\varepsilon)_{t\geqslant 0}$ (page 52)
$(Z_{s,t}^\xi)_{t\geqslant 0}$	small jump process belonging to $(Z_t^\varepsilon)_{t\geqslant 0}$ (page 61)
$(\hat{X}_t^\varepsilon)_{t\geqslant 0}$	time-scaled version of $(X_t^\varepsilon)_{t\geqslant 0}$ (page 63)
$(\hat{Z}_{s,t}^\varepsilon)_{t\geqslant 0}$	time-scaled version of $(Z_t^\varepsilon)_{t\geqslant 0}$ (page 73)
$\lambda(\varepsilon)$	characteristic time scale (page 92)
$(Y_t^\varepsilon)_{t\geqslant 0}$	jump diffusion with V as drift subject to a periodic additive process (page 76)
$(\widetilde{Y}_{s,t}^\varepsilon)_{t\geqslant 0}$	small jump process belonging to $(Y_t^\varepsilon)_{t\geqslant 0}$ (page 83)

Regions of attraction and the separatrix for the equation $\dot{x}(t) = -\nabla U(x(t), \frac{t}{2T})$:

$\Omega_\pm(t_0)$	domain of attraction of $p_\pm(\cdot)$ (page 50)
$\Gamma(t_0)$	separatrix belonging to $\Omega_\pm(t_0)$ (page 50)
$\Omega_\pm^{\varepsilon^\gamma}(t_0), \Omega_{\pm,R}^{\varepsilon^\gamma}(t_0)$, $\Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0), \Omega_{\pm,R}^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0)$, $\Omega_\pm^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t_0), \Omega_{\pm,R}^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t_0)$	reduced respectively reduced and bounded domains of attraction of the periodic solution $p_\pm(\cdot)$ (page 51)
$\Gamma^{\varepsilon^\gamma}(t_0), \Gamma_R^{\varepsilon^\gamma}(t_0)$, $\Gamma^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0), \Gamma_R^{\varepsilon^\gamma,\varepsilon^{2\gamma}}(t_0)$ $\Gamma^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t_0), \Gamma_R^{\varepsilon^\gamma,\varepsilon^{2\gamma},\varepsilon^{2\gamma}}(t_0)$	enlarged respectively enlarged and truncated separatrices (page 51)

Regions of attraction and the separatrix for the equation $\dot{y}(t) = -\nabla V(y(t))$:

$\Omega_\pm$	domains of attraction of minima $m_\pm$ of V (page 81)
$\Omega_\pm(\delta_1)$, $\Omega_{\pm,L}(\delta_1)$, $\Omega_\pm(\delta_1,\delta_2)$, $\Omega_{\pm,L}(\delta_1,\delta_2)$, $\Omega_\pm(\delta_1,\delta_2,\delta_3)$, $\Omega_{\pm,L}(\delta_1,\delta_2,\delta_3)$	reduced respectively reduced and bounded domains of attraction of minima of V (page 82)
$\Gamma(\delta_1)$, $\Gamma_L(\delta_1)$, $\Gamma(\delta_1,\delta_2)$, $\Gamma_L(\delta_1,\delta_2)$, $\Gamma(\delta_1,\delta_2,\delta_2)$, $\Gamma_L(\delta_1,\delta_2,\delta_2)$	enlarged respectively enlarged and truncated separatrices (page 82)

Exit times :

$\hat{\tau}_\pm^\varepsilon$	exit time of $(\hat{X}_t^\varepsilon)_{t\geqslant 0}$ from $\Omega_{\pm,R}^{\varepsilon^\gamma}(2Tt)$ (page 63)
$\hat{\tau}_\pm^{\varepsilon,M}$	exit time of $(\hat{Z}_t^\varepsilon)_{t\geqslant 0}$ from $\Omega_{\pm,R}^{\varepsilon^\gamma}(2Tt)$ (page 73)
$\tau_\pm^{\varepsilon,*}$	exit time of $(Y_t^\varepsilon)_{t\geqslant 0}$ from $\Omega_{\pm,L}(\delta_1^*)$ (page 92)

Ehrenwörtliche Erklärung

Hiermit erkläre ich,

- dass mir die Promotionsordnung der Fakultät bekannt ist,
- dass ich die Dissertation selbst angefertigt habe, keine Textabschnitte oder Ergebnisse eines Dritten oder eigene Prüfungsarbeiten ohne Kennzeichnung übernommen und alle von mir benutzten Hilfsmittel, persönliche Mitteilungen und Quellen in meiner Arbeit angegeben habe,
- dass ich die Hilfe eines Promotionsberaters nicht in Anspruch genommen habe und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,
- dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe.

Bei der Auswahl und Auswertung des Materials sowie der Herstellung des Manuskripts haben mich folgende Personen unterstützt:
Prof. Dr. Ilya Pavlyukevich, Friedrich-Schiller-Universität Jena

Ich habe die gleiche, eine in wesentlichen Teilen ähnliche bzw. eine andere Abhandlung bereits bei einer anderen Hochschule als Dissertation eingereicht: Ja / Nein
Wenn Ja, Name der Hochschule:
Ergebnis:

Isabelle Kuhwald | Jena, den 22.06.2015

Lebenslauf

Name:	Kuhwald
Vorname:	Isabelle
Geburtsdatum:	21.09.1986
Geburtsort:	Jena
Familienstand:	verheiratet
Kinder:	eine Tochter (geboren am 14.12.2012)

Schulbildung:

1997 - 2005	Staatliches Gymnasium "Ernst Haeckel" in Jena, Abschluss: Abitur (Note 1.1)

Hochschulbildung:

10/2005 - 11/2009	Studium der Mathematik an der Friedrich-Schiller-Universität Jena, Abschluss: Diplom (Gesamtprädikat: Mit Auszeichnung), Vertiefungsrichtung Stochastik/Statistik, Diplomarbeit mit dem Thema "The Brownian sheet - Existence and properties"

Berufliche Tätigkeit:

01/2010 - 06/2015	Wissenschaftliche Mitarbeiterin am Institut für Mathematik der Friedrich-Schiller-Universität Jena
11/2012 - 12/2013	Unterbrechnung der Arbeit als wissenschaftliche Mitarbeiterin durch Mutterschutz und Elternzeit

Jena, den 22.06.2015